EXISTENCE AND DIFFERENTIABILITY THEOREMS FOR THE SOLUTIONS OF VARIATIONAL PROBLEMS FOR MULTIPLE INTEGRALS[1]

CHARLES B. MORREY, JR.

I. Introduction

1. The object of the research [1] with which this address is concerned is twofold: first, to demonstrate, by direct methods, the existence of solutions, perhaps in some generalized sense, for a wide class of variational problems for multiple integrals, and second, to investigate further differentiability properties of the generalized solutions thus obtained. We consider only problems in nonparametric form.

2. Both of these goals have been achieved in the recent well known solutions of the problem of Plateau and in the solutions of the Dirichlet problem by variational methods. In more general problems this program has not been carried through completely except in very restricted cases. However, there are very important known results in connection with each separate aim. In connection with the existence theory, very important work has been done in the nonparametric case by Tonelli [2] and in the parametric case by McShane [3]. In the parametric case, practically nothing is known concerning the differentiability of the solutions obtained. In the nonparametric case, it has been proved by E. Hopf [4] that if $f(x, y, z, p, q)$ is of class C_α'' (that is, if f is of class C'' and its second derivatives satisfy a uniform Hölder condition[2] with exponent α, $0<\alpha<1$, on any bounded portion of space) and is the integrand in a regular variational problem (that is, if $f_{pp}f_{qq}-f_{pq}^2>0$, $f_{pp}>0$) and if z_0 is continuous on $\overline{G}$ and is of class C_β' on each region D with $\overline{D}\subset G$ and minimizes $\iint_G f(x, y, z, p, q)dxdy$ among all functions, continuous on $\overline{G}$, of class C' in G, and coinciding with z_0 on G^*, then z_0 is of class C_γ'' on any region D as above for any $\gamma<\alpha$. The author has proved in a previous paper [5] that the same result holds if z_0 merely satisfies a uniform Lipschitz condition[3] on regions $\overline{D}$ interior to G. The result of E. Hopf

[1] An address delivered before the meeting of the Society in Pasadena on December 2, 1939, by invitation of the Program Committee.

[2] A function $f(P)$ is said to satisfy a *uniform Hölder condition* on a set S if $|f(P_1)-f(P_2)|\leq C\cdot|P_1P_2|^\alpha$, $0<\alpha<1$, for each pair of points (P_1, P_2) in S; C is called the *coefficient* and α is called the *exponent* of the Hölder condition and $|P_1P_2|$ denotes the distance from P_1 to P_2; C and α are supposed to be independent of P_1 and P_2.

[3] A function $f(P)$ is said to satisfy a uniform Lipschitz condition on a set S if $|f(P_1)-f(P_2)|\leq C\cdot|P_1P_2|$ for every pair (P_1, P_2) on S, C independent of P_1, P_2.

can be extended readily to the case where f is a function of $(x, y, z, \cdots, z_N)$ while the methods used in the author's previous paper [5] do not permit this extension.

3. In connection with the existence theory, Tonelli proves in the two papers mentioned that if $f(x, y, z, p, q)$ satisfies certain continuity restrictions, gives rise to a quasi-regular variational problem, and if it satisfies

$$f(x, y, z, p, q) \geqq m(p^2 + q^2)^{\alpha/2} + N, \quad m > 0, \alpha > 2, \tag{3.1}$$

or (essentially)

$$f(x, y, z, p, q) \geqq m(p^2 + q^2), \quad f(x, y, z, 0, 0) = 0, \quad m > 0, \tag{3.2}$$

then the integral

$$I(z, G) = \iint_G f(x, y, z, p, q)dxdy$$

takes on its minimum among all functions z which are absolutely continuous in his sense and assume given continuous boundary values, provided $I(z, G)$ is finite for some such function and G is properly restricted; more general boundary value problems are also considered. The proof consists in showing first that $I(z, G)$ is lower semi-continuous with respect to uniform convergence in his class, and then in replacing any minimizing sequence by one in which the functions are equicontinuous so that a uniformly convergent minimizing subsequence is obtained which can be shown to converge to a function in his class. If the α in (3.1) is less than 2, such a procedure is not possible in general. In fact, when we try to generalize these results to functions of N independent variables, we find that Tonelli's procedure is impossible unless $\alpha \geqq N$ (see §12). As the gap $1 < \alpha < N$ grows wider as N increases, it seems desirable to obtain existence theorems of some sort assuming merely that $\alpha > 1$; for $\alpha = 1$, there are examples, even for $N = 1$, of problems which have no solutions.

Evidently we must abandon the use of continuity and equicontinuity in obtaining our existence theorems but must certainly retain absolute continuity in some sense. G. C. Evans [6] has introduced a class of functions, which he calls "potential functions of their generalized derivatives," which would seem to form a proper class of admissible functions; we shall call these functions "functions of class $\mathfrak{P}$" (see below). Surprisingly enough, these functions were introduced before those of Tonelli, and Evans [7] has investigated the connection between the two classes of functions (we state the results below).

4. Before proceeding to a discussion of these functions and their applications, we must introduce certain notations and terminology. The N-tuples $(x_1, \cdots, x_N)$, $(\xi_1, \cdots, \xi_N)$, $\cdots$, will be denoted by the single letters x, ξ, $\cdots$, and the closed cell $a_i \leqq x_i \leqq b_i$, $i=1, \cdots, N$, will be denoted by $[a, b]$. The functional notation $f(x_1, \cdots, x_N)$ will be abbreviated to $f(x)$ and the Lebesgue integrals of $f(x)$ over $[a, b]$ or a set E will be denoted by

$$\int_a^b f(x)dx, \qquad \int_E f(x)dx.$$

It is frequently desirable to consider the behavior of a function $f(x)$ with respect to a particular coordinate variable x_k or with respect to the $(N-1)$ variables $(x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_N)$. We then denote the latter variables by x_k', we write (x_k', x_k) for x and $f(x_k', x_k)$ for $f(x)$: thus $f(a_k', x_k)$ denotes the function $f(a_1, \cdots, a_{k-1}, x_k, a_{k+1}, \cdots, a_N)$ of the single variable x_k (a_k' being fixed) and $f(x_k', a_k)$ denotes the function $f(x_1, \cdots, x_{k-1}, a_k, x_{k+1}, \cdots, x_N)$ (a_k being fixed). The projection of the cell $[a, b]$ on $x_k=0$ is denoted by $[a_k', b_k']$ and

$$\int_{a_k'}^{b_k'} f(x_k', a_k)dx_k'$$

denotes the $(N-1)$-dimensional integral of $f(x_k', a_k)$ over $[a_k', b_k']$. If G is a region (open connected set), G^* denotes its boundary and $\overline{G}$ its closure. All integrals are Lebesgue integrals.

II. The admissible functions

5. We now define the various classes of functions of which we shall speak.

DEFINITION 5.1. *A function $f(x)$ is of class $\mathfrak{P}$ on a region G if*

(i) *it is summable on each closed cell interior to G, and*

(ii) *there exist functions $v_1(x), \cdots, v_N(x)$, satisfying* (i) *and such that*

$$\int_{a_k'}^{b_k'} [f(x_k', b_k) - f(x_k', a_k)]dx_k' = \int_a^b v_k(x)dx$$

for almost all cells $[a, b]$ interior to G (that is, all such cells such that the point $(a_1, \cdots, a_N; b_1, \cdots, b_N)$ does not belong to a set of $2N$-dimensional measure zero).

DEFINITION 5.2. *A function $f(x)$ is of class $\mathfrak{P}'$ on G if*

(i) *it satisfies* (i) *of Definition* 5.1, *and*

(ii) *if* $[a, b]$ *is any cell interior to* G, $f(x_k', x_k)$ *is A.C. (absolutely continuous) in* x_k *for* $a_k \leqq x_k \leqq b_k$ *for almost every* x_k' *in* $[a_k', b_k']$, *and*

(iii) $\partial f/\partial x_k$ *(which exists almost everywhere and is measurable by* (i) *and* (ii)) *is summable on each cell interior to* G.

DEFINITION 5.3 [8]. *A function* $f(x)$ *is absolutely continuous in the sense of Tonelli (A.C.T.) on* G *if*

(i) *it is continuous on* G,

(ii) *it satisfies* (ii) *of Definition* 5.2, *and*

(iii) *if* $[a, b]$ *is any cell in* G, *then the variation of* $f(x_k', x_k)$ *as a function of* x_k *alone on* $[a_k, b_k]$ *is a summable (known to be lower semicontinuous from* (i)) *function of* x_k'.

DEFINITION 5.4. *A function* $f(x)$ *is of class* $\mathfrak{P}''$ *on* G *if it is of class* $\mathfrak{P}$ *there and is continuous.*

It is easily shown that a function of $f(x)$ of class $\mathfrak{P}$ determines the functions $v_k(x)$ of Definition 5.1 to within a null function (a function which vanishes almost everywhere). Hence it defines completely the set functions $\int_e v_k(x)dx$ $(k=1, \cdots, N)$ which are A.C. on any bounded set E whose closure is in G. Accordingly we propose, with Evans [6], the following definition:

DEFINITION 5.5. *The generalized derivative* $D_{x_k}f$ *of* f *with respect to* x_k *is defined as the Lebesgue derivative* [9] *of the set function* $\int_e v_k(x)dx$ *defined above, wherever this derivative exists.*

It is well known that this derivative exists almost everywhere and coincides with $v_k(x)$ almost everywhere. For simplicity, we shall use the notation f_{x_k} to denote either $D_{x_k}f$ or $\partial f/\partial x_k$ as desired (see §6 for a justification of this usage).

6. The following theorems concerning the above classes of functions are easily demonstrated:

THEOREM 6.1. *If* $f(x)$ *is of class* $\mathfrak{P}$ *on a region* G, *any function* $f^*(x)$ *which is equivalent to* $f(x)$ *on* G *(that is, differs from* $f(x)$ *by a null function) is also of class* $\mathfrak{P}$ *on* G *and the generalized derivatives of* f *and* f^* *exist at exactly the same points and coincide at these points.*

THEOREM 6.2. *Any two functions of class* $\mathfrak{P}$ *whose respective generalized derivatives coincide almost everywhere differ at most by a constant and a null function.*

THEOREM 6.3. *Any function of class* $\mathfrak{P}'$ *is of class* $\mathfrak{P}$ *and corresponding generalized and partial derivatives coincide almost everywhere.*

THEOREM 6.4. *Any function of class* $\mathfrak{P}$ *is equivalent to a function of class* $\mathfrak{P}'$.

THEOREM 6.5. *The classes in Definitions* 5.3 *and* 5.4 *coincide.*

THEOREM 6.6. *A necessary and sufficient condition that* $f(x)$ *be of class* $\mathfrak{P}$ *on* G *is that it satisfy* (i) *of Definition* 5.1 *and that for each cell* $[a, b]$ *interior to* G, *there exist summable functions* $v_k(x)$ *and a sequence* $\{f_n(x)\}$, *each* f_n *satisfying a uniform Lipschitz condition on* $[a, b]$, *such that*

$$\lim_{n\to\infty} \int_a^b \left[|f_n - f| + \sum_{k=1}^{N} |f_{n,x_k} - v_k| \right] dx = 0.$$

7. We now define the most important of all of our classes of admissible functions. The theorem below and those in the next five sections would make it appear that these functions should play an important rôle in many other branches of analysis as well as in the present research.

DEFINITION 7.1. *A function* $f(x)$ *is said to be of class* $\mathfrak{P}_\alpha$ ($\mathfrak{P}_\alpha'$ *or* $\mathfrak{P}_\alpha''$), $\alpha \geqq 1$, *on* G *if it is of class* $\mathfrak{P}$ ($\mathfrak{P}'$ *or* $\mathfrak{P}''$) *there and if* $|f|^\alpha$ *and* $|D_{x_k} f|^\alpha$ *are summable over* G, $k = 1, \cdots, N$.

The following theorem concerning these functions is of fundamental importance.

THEOREM 7.1. *The space* $\mathfrak{P}_\alpha$ *of classes of equivalent functions of class* $\mathfrak{P}_\alpha$ *on a region* G *is a Banach space* (*space of type B*) [10] *if we define, for instance,*

$$\|f\|^\alpha = \overline{D}_\alpha(f, G) \equiv D_\alpha(f, G) + \int_G |f|^\alpha dx,$$

(7.1)

$$D_\alpha(f, G) \equiv \int_G \left[\sum_{i=1}^{N} |f_{x_i}|^2 \right]^{\alpha/2} dx,$$

$\|f\|$ *denoting the norm of* f *in* $\mathfrak{P}_\alpha$ *on* G.

8. Before proceeding further with the discussion of these functions, we propose two more definitions as follows:

DEFINITION 8.1. *A transformation* $x = x(y)$ *of a set* T *into a set* S *is Lipschitzian if it is* 1-1 *and continuous and if the functions* $x_i(y)$ *and* $y_i(x)$ (*of the inverse*) *all satisfy uniform Lipschitz conditions on each closed subset of* T *and* S *respectively. If the Lipschitz conditions are uniform over the whole of* T *and* S, *we say that the Lipschitzian transformation is regular.*

DEFINITION 8.2. *A region G is said to be Lipschitzian if $\overline{G}$ can be covered by a finite number of sets $\gamma_1, \cdots, \gamma_L$, each of which is the product of $\overline{G}$ with an open set and is the image under a regular Lipschitzian transformation of either the open unit cube $|y_i| < 1$ or the part of this cube for which $y_N \leqq 0$; in the latter case we assume that the transformation sets up a 1-1 correspondence between the points of $\gamma_i \cdot G^*$ and the points of the half-cube for which $y_N = 0$.*

9. Rademacher [11] has shown that a function satisfying a Lipschitz condition on a region G possesses a total differential almost everywhere and has also proved the usual theorems on transformations of multiple integrals to hold for Lipschitzian transformations. With these definitions and results in mind, we may generalize a result of Evans [12] as follows:

THEOREM 9.1. *Let $f(x)$ be of class $\mathfrak{P}$ (or $\mathfrak{P}''$) on a region G and let $x = x(y)$ be a Lipschitzian transformation of a region H into G. Then the transformed function $f[x(y)] = g(y)$ is of class $\mathfrak{P}$ (or $\mathfrak{P}''$) on H. Moreover if $x_0 = x(y_0)$ where y_0 is a point of H at which all the $x_i(y)$ possess total differentials and if all the generalized derivatives of f exist at x_0, then all the generalized derivatives of g exist at y_0 and are connected with those of f at x_0 by the usual formulas. If f is of class $\mathfrak{P}_\alpha$ on G and the transformation is regular, then g is of class $\mathfrak{P}_\alpha$ on H.*

Concerning the functions of class $\mathfrak{P}'$, we have proved the following theorem:

THEOREM 9.2. *Any function of class $\mathfrak{P}$ on G is equivalent to a function of class $\mathfrak{P}'$ on G which is transformed into a function of class $\mathfrak{P}'$ by any regular Lipschitzian transformation.*

Not every function of class $\mathfrak{P}'$ has this property, and even if it does, an example of Saks [13] shows that the statement of Evans concerning the generalized derivatives is not in general true for the partial derivatives.

10. We now come to a discussion of the boundary values of functions of class $\mathfrak{P}_\alpha$ on bounded regions. If $f(x)$ is of class $\mathfrak{P}_\alpha$ on a Lipschitzian region G ($\alpha \geqq 1$), we may state the following fundamental results:

THEOREM 10.1. *There exists a sequence $\{f_n(x)\}$ of functions, each satisfying a uniform Lipschitz condition on $\overline{G}$, which converges strongly, accordingly to the norm (7.1) to f on G.*

THEOREM 10.2. *There exists a boundary value function f^* of class L_α*

on G^* (with respect to hyper-area on G^*) such that f_n converges strongly in L_α on G^* to f^*, $\{f_n\}$ being any sequence as in Theorem 10.1.

THEOREM 10.3. *If $f^*=0$ almost everywhere, the sequence $\{f_n(x)\}$ of Theorem* 10.1 *may be chosen so that each f_n vanishes on and near G^*.*

THEOREM 10.4. *If $x'=x'(x)$ is a regular Lipschitzian transformation of G into G' and if f' and f'^* are the transforms of f and f^*, then f'^* is the boundary value function for f' on G'^*.*

THEOREM 10.5. *Let $\bar{f}(x)$ be of class $\mathfrak{P}_\alpha'$ and equivalent to $f(x)$ on G. Then there exist sets Z_k, $k=1, \cdots, N$, each of $(N-1)$-dimensional measure zero, such that if $\bar{x}_k'$ is not in Z_k, $\bar{f}(\bar{x}_k', x_k)$ is A.C. on each segment $x_k^{(1)} \leqq x_k \leqq x_k^{(2)}$, such that $(\bar{x}_k', x_k)$ is in G and $(\bar{x}_k', x_k^{(i)})$ is on G^*, $i=1, 2$, and $\bar{f}(\bar{x}_k', x_k)$ tends to $f^*(\bar{x}_k', x_k^{(i)})$ as x_k tends to $x_k^{(i)}$ along this segment, $i=1, 2$, $k=1, \cdots, N$.*

If G is merely an arbitrary bounded region, we do not have such explicit results but we may follow a procedure suggested by Courant [14]. This procedure suggests the following definition:

DEFINITION 10.1. *A function $f(x)$ of class $\mathfrak{P}_\alpha$ on G is said to vanish on G^* if there exists a sequence $\{f_n(x)\}$ of functions, each satisfying a uniform Lipschitz condition on $\overline{G}$ and vanishing on and near G^*, such that $\|f_n-f\| \to 0$. Two functions f_1 and f_2 of class $\mathfrak{P}_\alpha$ are said to coincide on G^* if their difference vanishes on G^* in the above sense.*

From Theorem 10.3, it is clear that the above definition of vanishing coincides with the condition that $f^*=0$ in the case that G is Lipschitzian. Using this very general terminology, we may state the following very important "substitution theorem":

THEOREM 10.6. *Let $f(x)$ be of class $\mathfrak{P}_\alpha$ on a region G, let $g(x)$ be of class $\mathfrak{P}_\alpha$ on a subregion D of G and coincide with $f(x)$ on D^*, and let $h(x)$ be defined equal to f on $G-D$ and equal to g on D. Then $h(x)$ is of class $\mathfrak{P}_\alpha$ on G, coincides with $f(x)$ on G^*, and $D_{x_k}h=D_{x_k}f$ almost everywhere on $G-D$ (D is open) and $D_{x_k}h=D_{x_k}g$ on D wherever either exists, $k=1, \cdots, N$.*

11. As the spaces $\mathfrak{P}_\alpha$ have been seen to be Banach spaces, we may inquire into the nature and consequences of weak convergence [15] in $\mathfrak{P}_\alpha$. In this connection, the following interesting results have been proved:

THEOREM 11.1. *The most general linear functional* [16] *$F(f)$ defined on $\mathfrak{P}_\alpha$ is of the form*

$$F(f) = \int_G \left[A_0(x)f(x) + \sum_{i=1}^{N} A_i(x) D_{x_i} f \right] dx,$$

where the A_i are essentially bounded if $\alpha = 1$ and are in L_β with $\alpha^{-1} + \beta^{-1} = 1$ if $\alpha > 1$ (the $A_i(x)$ are not uniquely determined by $F(f)$ however).

THEOREM 11.2. *A necessary and sufficient condition that f_n tend weakly to f in $\mathfrak{P}_\alpha$ is that f_n and $D_{x_i} f_n$ tend weakly to f and $D_{x_i} f$, respectively.*

THEOREM 11.3. *Weak convergence is preserved by regular Lipschitzian transformations.*

THEOREM 11.4. *If f_n tends weakly to f in G, the convergence holds on any subregion and f is uniquely determined.*

THEOREM 11.5. *If the f_n all vanish on G^*, f vanishes on G^*.*

THEOREM 11.6. *If G is Lipschitzian and f_n tends weakly in $\mathfrak{P}_\alpha$ on G to f, then f_n tends strongly in L_α to f on G and the boundary value functions f_n^* tend strongly in L_α on G^* to f^*.*

THEOREM 11.7. *If $\alpha > 1$, a necessary and sufficient condition that a family of functions $\{f\}$ in $\mathfrak{P}_\alpha$ be compact (perhaps not closed) with respect to weak convergence in $\mathfrak{P}_\alpha$ is that their norms be uniformly bounded. If $\alpha = 1$ we must also have a convex function $\phi(p_1, \cdots, p_N)$ with*

$$\lim_{|p| \to \infty} \phi(p_1, \cdots, p_N)/|p| = +\infty, \qquad |p|^2 = p_1^2 + \cdots + p_N^2,$$

such that

$$\int_G \phi[D_{x_1} f, \cdots, D_{x_N} f] dx$$

is uniformly bounded.

THEOREM 11.8. *A necessary and sufficient condition that the norms in $\mathfrak{P}_\alpha$ of each function of a family be uniformly bounded is that $D_\alpha(f, G)$ be uniformly bounded, each function f of the family coinciding on G^* with a function g of another family whose norms are uniformly bounded.*

THEOREM 11.9. *If G is Lipschitzian, the above condition may be replaced by either the condition that $\overline{D}_\alpha(f, \tau)$ be uniformly bounded for some cell τ in G or the condition that $\int_E |f^*|^\alpha ds(e)$ be uniformly bounded, E being some set open on G^* and $s(e)$ being the hyper-area set function on G^*.*

12. Of less interest here are Green's theorem for functions of class $\mathfrak{P}_1$ on a Lipschitzian region G and theorems hinted at above (in §3) which show that functions of class $\mathfrak{P}_\alpha$ with $\alpha > N$ are equivalent to continuous functions. If $\alpha \leqq N$, then this is not usually true as the examples

$$f(x) = \log\log(1 + r^{-1}), \qquad \alpha = N,\ r^2 = x_1^2 + \cdots + x_N^2,$$

$$f(x) = r^{-\beta}, \qquad \alpha < N,\ 0 < \beta,\ \alpha(\beta + 1) < N,$$

show.

III. The existence theory

13. Having studied our admissible functions at some length, we can now easily derive some very general existence theorems. We consider integrals

$$I(z, G) = \int_G f(x_1, \cdots, x_N, z_1, \cdots, z_P, p_1^1, \cdots, p_N^1, \cdots, p_1^P, \cdots, p_N^P)dx$$

$$= \int_G f(x, z, p)dx, \qquad p_j^i = D_{x_j} z_i,$$

in which $f(x, z, p)$ is continuous all over the space of its arguments, is convex in p for each (x, z), and satisfies a condition

$$f(x, z, p) \geqq \phi(p), \qquad \lim_{|p| \to \infty} |p|^{-1}\phi(p) = +\infty, \qquad |p|^2 = \sum_{i=1}^{N}\sum_{j=1}^{P} (p_i^j)^2,$$

$\phi(p)$ being convex. It is straightforward to prove that such an integral $I(z, G)$ is lower semicontinuous with respect to weak convergence in $\mathfrak{P}_1$. From our study of the admissible functions, we see that any family of functions z for which $I(z, G)$ is uniformly bounded will be compact with respect to weak convergence in $\mathfrak{P}_1$, provided merely that the norms are uniformly bounded. This may be ensured by satisfying one of the conditions in Theorems 11.8 or 11.9. If this can all be done for a minimizing sequence, we may then pick out a subsequence which converges weakly in $\mathfrak{P}_1$ to some function *which is also in* $\mathfrak{P}_1$. From the lower semicontinuity of $I(z, G)$ we may conclude that our limit function is a minimizing function. It is clear that a great variety of existence theorems with variable as well as fixed boundary values can be proved with very little difficulty. For example, we may conclude the existence of a minimal surface part of whose boundary is to be a fixed Jordan arc or is to lie on a bounded closed manifold. Of course our existence theorems do not allow us to conclude continuity

of the surface on the boundary and, indeed, this is quite difficult to prove and may not be true in many problems.

IV. Continuity properties of the minimizing functions

14. Without placing further restrictions on the integrand $f(x, z, p)$, one cannot expect to be able to deduce many further properties of the solutions obtained from the above existence theory. This type of study seems quite difficult and is almost unexplored. Accordingly we have restricted ourselves to the case $N=2$ and have required that our function (N is the previous P)

$$f(x, y, z^1, \cdots, z^N, p^1, \cdots, p^N, q^1, \cdots, q^N) = f(x, y, z, p, q)$$

be of class C_a'' over the space of its arguments and that it satisfy the further conditions

$$\begin{aligned} &m_1(p^2+q^2) \leqq f(x, y, z, p, q) \leqq M_1(p^2+q^2), \\ &p^2+q^2 = \sum_{i=1}^{N} [(p^i)^2 + (q^i)^2], \quad 0 < m_1 \leqq M_1, \end{aligned} \tag{14.1}$$

$$\begin{aligned} &m_2(R)(\xi^2+\eta^2) \leqq f_{p^\alpha p^\beta}\xi^\alpha\xi^\beta + 2f_{p^\alpha q^\beta}\xi^\alpha\eta^\beta + f_{q^\alpha q^\beta}\eta^\alpha\eta^\beta \\ &\qquad \leqq M_2(R)(\xi^2+\eta^2), \quad \xi^2+\eta^2 = \sum_{i=1}^{N} [(\xi^i)^2 + (\eta^i)^2], \\ &\sum_{i=1}^{N}\Big[f^2_{p^i x} + f^2_{p^i y} + f^2_{q^i x} + f^2_{q^i y} + |f_{z^i x}| + |f_{z^i y}| \\ &\qquad + \sum_{j=1}^{N}\{f^2_{p^i z^j} + f^2_{q^i z^j} + |f_{z^i z^j}|\}\Big] \leqq M_3(R)(p^2+q^2), \\ &0 < m_2(R) \leqq M_2(R), \quad x^2+y^2+(z^1)^2+\cdots+(z^N)^2 \leqq R^2, \end{aligned} \tag{14.2}$$

where in (14.1) we assume m_1 and M_1 to be independent of (x, y, z, p, q) and where $m_2(R)$, $M_2(R)$, and $M_3(R)$ depend only on R.

This type of integrand, although quite restricted, is not trivial. For instance, the integrand obtained in the problem of Plateau on a Riemannian manifold (if Courant's method [17] for the ordinary problem of Plateau is used) is

$$g_{\alpha\beta}(z^1, \cdots, z^N)(p^\alpha p^\beta + q^\alpha q^\beta), \tag{14.3}$$

which satisfies all the conditions (14.1) and (14.2) on the cell of the z space in which the $g_{ij}(z)$ are defined, provided merely that the part of the manifold thus represented is regular and the g_{ij} are of class C_a''. Moreover, the existence theory for such an integrand does not follow

from the results of Tonelli as the corresponding restrictions (3.2) for his methods to be extended to N dependent variables $z^1, \cdots, z^N$ would be (essentially)

$$f(x, y, z, p, q) \geqq m(p^2 + q^2), \qquad m > 0,$$

$$f(x, y, z^1, \cdots, z^N, p^1, \cdots, p^{k-1}, 0, p^{k+1}, \cdots, p^N, q^1, \cdots, q^{k-1}, 0, q^{k+1}, \cdots, q^N) = 0, \quad k = 1, \cdots, N,$$

such a set of equalities not being assumed in our case. In fact it seems very difficult to show the existence of solutions using methods which involve equicontinuity.

We have used the tensor summation convention in this section, summing only over repeated Greek indices, and we shall continue its use throughout the rest of the paper.

15. Of fundamental importance in the continuity and differentiability theory are the following four theorems:

THEOREM 15.1. *If the vector function z (that is, each component z^i, $i=1, \cdots, N$) is of class $\mathfrak{P}_2$ on G and if*

$$D_2[z, C(P, r)] \leqq M(r/a)^\lambda \tag{15.1}$$

for each circle $C(P, r)$ in G, then z (is equivalent to a function which) satisfies a uniform Hölder condition with exponent $\frac{1}{2}\lambda$ on each region D with $\overline{D} \subset G$, the coefficient in which depends only on M, λ, and the distance of $\overline{D}$ from G^*; a denotes the distance of P from G^*.

THEOREM 15.2. *If, in Theorem* 15.1, (15.1) *holds only for circles with center at a fixed point P, then the average of z over circles $C(P, r)$ tends to a limit $\bar{z}(P)$ as $r \to 0$. If z vanishes on G^*, then*

$$|\bar{z}(P)| \leqq J(M, \lambda, a, \delta), \tag{15.2}$$

where J is a number depending only on the indicated quantities, δ being the diameter of G and a being the distance of P from G^*.

THEOREM 15.3. *Let z be of class $\mathfrak{P}_2$ on G and satisfy*

$$D_2(z, \Delta) \leqq K \cdot D_2[H(z, \Delta), \Delta], \qquad K \geqq 1, \tag{15.3}$$

for each region Δ in G, $H(z, \Delta)$ being the harmonic function in Δ coinciding with z on Δ^* (known to exist by our existence theorems and easily proved harmonic). Suppose G is bounded by a finite number of nonintersecting simple closed curves and suppose that z coincides with a function of class $\mathfrak{P}_2'$ on G and continuous on $\overline{G}$.

Then z is continuous on $\overline{G}$ (or equivalent to such a function).

THEOREM 15.4. *If z is of class $\mathfrak{P}_2$ on G and satisfies*

$$D_2[z, C(P, r)] \leqq K \cdot D_2[H\{z, C(P, r)\}, C(P, r)] + L \cdot (r/a)^{\lambda}, \quad 0 < \lambda < K^{-1}, \tag{15.4}$$

for all circles in G with center at a fixed point P $(0 \leqq r \leqq a)$, then z satisfies (15.1), *the M of* (15.1) *depending only on K, L, and λ.*

Some of these theorems have generalizations to vector functions of n $(n \geqq 2)$ variables. To generalize Theorems 15.1 and 15.2, we need only to replace the λ of (15.1) by $n-2+\lambda$, $\lambda > 0$. Theorem 15.3 has no immediate generalization for all numbers K and Theorem 15.4 yields an exponent $(n-1)/K$ which is not greater than $n-2$ unless K is restricted to be quite near to unity. As we use Theorem 15.4 to obtain a condition (15.1), it will thus appear to the reader that the methods used in this research will not generalize immediately to integrands of the type described in §14 where x and y are replaced by n independent variables, essential use being made of the bound (15.2) and the continuity results of Theorem 15.1, both of which require the exponent $n-2+\lambda$, $\lambda > 0$.

16. Now, if z minimizes our integral $I(z, G)$ among all functions of class $\mathfrak{P}_2$ which coincide with it on G^*, then it follows from (14.1) that z satisfies the condition in Theorem 15.3 with $K = M_1/m_1$; for if some D exists where this does not hold, we can define $Z(x) = z(x)$ on $G-D$ and define $Z(x) = H(z, D)$ on D; then $Z(x)$ is of class $\mathfrak{P}_2$ on G and we have (using Theorem 10.6)

$$\begin{aligned} I(Z, G) &= I(z, G - D) + I[H(z, D), D] \\ &\leqq I(z, G - D) + M_1 D_2[H(z, D), D] \\ &< I(z, G - D) + m_1 D_2(z, D) \leqq I(z, G), \end{aligned} \tag{16.1}$$

so that z would not minimize $I(z, G)$. Thus, from the theorems in §15, we see that $D_2[z, C(P, r)] \leqq (r/a)^{\lambda} \cdot D_2(z, G)$, $\lambda = m_1/M_1$, for each circle in G with $r \leqq a$, that z satisfies a uniform Hölder condition, with exponent $\frac{1}{2}\lambda$ on closed regions interior to G, and if the boundary values are continuous and G is bounded by a finite number of simple closed curves, then z is continuous on $\overline{G}$.

17. To proceed with the examination of our minimizing function z, we next can extend Haar's well known lemma [18] to show that z satisfies

$$\int_{R^*} (f_p{}^i dy - f_q{}^i dx) = \iint_R f_z{}^i dx dy \tag{17.1}$$

for almost all rectangles R. If D is any simply connected subregion of G, we may find functions w^i and W^i of class $\mathfrak{P}_2'$ which satisfy

$$w^i_x - W^i_y = -f_{q^i}, \qquad w^i_y + W^i_x = f_{p^i},$$

$$\int_{R^*} (W^i_x dy - W^i_y dx) = \iint_R f_{z^i} dx dy, \tag{17.2}$$

almost everywhere and for almost all R on D. We also note that if D is any simply connected region with $\overline{D} \subset G$ and z is any minimizing function for $I(z, G)$ (whether continuous on the boundary or not), then (referring to (14.2) and various lemmas in potential theory) it follows that

$$\begin{aligned} D_2[z, D\cdot C(P, r)] &\leqq M_4 r^\lambda, \qquad D_2[w, D\cdot C(P, r)] \leqq M_4 r^\lambda, \\ D_2[W, D\cdot C(P, r)] &\leqq M_4 r^\lambda, \qquad \lambda > 0, \end{aligned} \tag{17.3}$$

and (14.2) holds with $R = \max_{z, y \in D} [x^2 + y^2 + (z^1)^2 + \cdots + (z^N)^2]$ which is bounded on such a region $\overline{D}$.

18. We next resort to a device due to Lichtenstein [19] and used by E. Hopf [4]. This consists in subtracting equation (17.1) formed for a rectangle $(a, c; b, d)$ $(a \leqq x \leqq b, c \leqq y \leqq d)$ from the same equation for the rectangle $(a+h, c; b+h, d)$ or the rectangle $(a, c+h; b, d+h)$ and then dividing by h. We confine ourselves to a region Δ such that $\overline{\Delta} \subset D \subset \overline{D} \subset G$. If we use the first rectangle above, we find that if h is fixed and sufficiently small,

$$\begin{aligned} &\int_{R^*} (a_{i\beta} u^\beta_x + b_{i\beta} u^\beta_y + d_{i\beta} u^\beta + g_i) dy \\ &\quad - (b_{\beta i} u^\beta_x + c_{i\beta} u^\beta_y + e_{i\beta} u^\beta + k_i) dx \\ &\quad = \iint_R (d_{\beta i} u^\beta_x + e_{\beta i} u^\beta_y + f_{i\beta} u^\beta + l_i) dx dy, \quad i = 1, \cdots, N, \end{aligned} \tag{18.1}$$

for almost all rectangles R in Δ where the coefficients are measurable functions and a_{ij}, for instance, is given almost everywhere by the formula

$$\begin{aligned} a_{ij}(x, y) = \int_0^1 f_{p^i p^j} [&x + th, y, z^i(x, y) + t\{z^i(x+h, y) - z^i(x, y)\}, \\ &p^i(x, y) + t\{p^i(x+h, y) - p^i(x, y)\}, \\ &q^i(x, y) + t\{q^i(x+h, y) - q^i(x, y)\}] dt \end{aligned} \tag{18.2}$$

and the other coefficients are given by similar formulas, and the $u^i(x, y)$ are the functions

$$u^i(x, y) = \frac{z^i(x + h, y) - z^i(x, y)}{h}. \tag{18.3}$$

It is easily seen, using (14.2), (17.3), and (18.2) that there exists an $h_0 > 0$ and numbers m, M, P, and μ with $0 < m \leqq M$ such that the coefficients satisfy (almost everywhere) the inequalities

$$\begin{gathered} m(\xi^2 + \eta^2) \leqq a_{\alpha\beta}\xi^\alpha\xi^\beta + 2b_{\alpha\beta}\xi^\alpha\eta^\beta + c_{\alpha\beta}\eta^\alpha\eta^\beta \leqq M(\xi^2 + \eta^2), \\ \iint_{\Delta \cdot C(P,r)} \sum_{i=1}^{N} \Big[|l_i| + \sum_{j=1}^{N} \{ d_{ij}^2 + e_{ij}^2 + |f_{ij}| \} \Big] dxdy \leqq Pr^\mu, \\ \iint_{\Delta \cdot C(P,r)} (g^2 + k^2) dxdy \leqq Pr^\mu, \qquad \mu > 0, \end{gathered} \tag{18.4}$$

for all circles $C(P, r)$ if $|h| < h_0$. By taking the rectangle $(a, c+h; b, d+h)$, we obtain similar equations and the same bounds as in (18.4) if h_0 is small enough, but in this case we have

$$u^i(x, y) = \frac{z^i(x, y + h) - z^i(x, y)}{h}. \tag{18.5}$$

We emphasize that the bounds m, M, P, and μ are independent of h if $|h| < h_0$. If D is simply connected, we can carry out simultaneously a similar process on the equations (17.2) and we get

$$\begin{gathered} v^i_x - V^i_y = -(b_{\beta i}u^\beta_x + c_{i\beta}u^\beta_y + e_{i\beta}u^\beta + k_i), \\ v^i_y + V^i_x = a_{i\beta}u^\beta_x + b_{i\beta}u^\beta_y + d_{i\beta}u^\beta + g_i, \\ \int_{R^*} V^i_x dy - V^i_y dx = \iint_R (d_{\beta i}u^\beta_x + e_{\beta i}u^\beta_y + f_{i\beta}u^\beta + l_i) dxdy, \\ i = 1, \cdots, N, \\ v^i(x, y) = \frac{w^i(x + h, y) - w^i(x, y)}{h}, \cdots, \\ V^i(x, y) = \frac{W^i(x + h, y) - W^i(x, y)}{h}, \cdots, \end{gathered} \tag{18.6}$$

almost everywhere and for almost all R. From (17.3) we have numbers Q and ν, independent of h if $|h| < h_0$ such that

$$(18.7) \qquad \iint_{\Delta \cdot C(P,r)} (u^2 + v^2 + V^2)dxdy \leqq Qr^\nu, \qquad \nu > 0,$$

for all $C(P, r)$.

19. In the course of this research, we have demonstrated the following theorem concerning the system (18.6):

THEOREM 19.1. *Any solution* (u, v, V) *of* (18.6) *which satisfies* (18.7) *on a region* Δ *is of class* $\mathfrak{P}_2''$ *and satisfies a uniform Hölder condition on each region* H *with* $\overline{H} \subset \Delta$; *the coefficient and exponent of this Hölder condition depend only on the diameter of* H, *the distance of* $\overline{H}$ *from* Δ^*, *and on the numbers* m, M, P, μ, Q, *and* ν *of* §18.

Thus if H is any region with $\overline{H} \subset G$, we may let $h \to 0$ in §18 and we then conclude (since we may then choose regions Δ and D with $\overline{H} \subset \Delta \subset \overline{\Delta} \subset D \subset \overline{D} \subset G$) that *the first derivatives of* z *satisfy uniform Hölder conditions on* $\overline{H}$. This result plus that of E. Hopf mentioned in §2 and those in §16 lead us to the final theorem:

THEOREM 19.2. *If the integrand* $f(x, y, z, p, q)$ *of* §§13 *and* 14 *satisfies the conditions of those two sections and if* z *minimizes* $I(z, G)$ *among all functions of class* $\mathfrak{P}_2$ *on a region* G *which coincide with* z *on* G^*, *then* z *is of class* C_β'' *on any region* H *with* $\overline{H} \subset G$ *for each* $\beta < \alpha$, α *being the* α *of* §14. *Moreover if* G *is bounded by a finite number of simple closed curves and if the boundary values of* z *are continuous on* G^*, *then* z *is continuous on* $\overline{G}$.

This result together with those of §13 completes the program announced in §1 for the important class of variational problems described in §14.

It should be remarked that S. Bochner [20] has been able to solve the *Euler differential equations* for the integrand (14.3) (probably by means of the topological methods of Leray and Schauder [21]) and to prove certain properties of the solutions. He has been unable, however, to demonstrate the minimizing property of his solutions. The advantage of the procedure in this paper for this purpose is evident. By using direct methods, the author [22] has recently solved the Plateau problem (one contour case) on a Riemannian manifold of considerable generality, the surface thus obtained then being shown, using the results in this paper, to be of class C''.

V. A STUDY OF THE SYSTEMS (18.1) AND (18.6)

20. The most interesting part of this research, however, is concerned with the study of the systems (18.1) and (18.6) (which is

equivalent to (18.1) if Δ is simply connected) which leads to the result stated in Theorem 19.1. Accordingly, we shall present an outline of this study.

It is easily seen that the equations (18.1), with $d_{ij}=e_{ij}=f_{ij}=l_i=0$ are the Haar equations for a function u which minimizes the integral

$$J(u,\Delta)=\iint_{\Delta}(a_{\alpha\beta}u_x^{\alpha}u_x^{\beta}+2b_{\alpha\beta}u_x^{\alpha}u_y^{\beta}+c_{\alpha\beta}u_y^{\alpha}u_y^{\beta}+2g_{\alpha}u_x^{\alpha}+2k_{\alpha}u_y^{\alpha}+H)dxdy$$

where H is chosen summable and just large enough to make the integrand non-negative. It is also easy to show that $J(u,\Delta)$ is lower semi-continuous with respect to weak convergence in $\mathfrak{P}_2$ and that minimizing functions exist which coincide on Δ^* with any given function of class $\mathfrak{P}_2$. Also if a function u_0 of class $\mathfrak{P}_2$ on Δ satisfies (18.1) with $d_{ij}=e_{ij}=f_{ij}=l_i=0$, then it can easily be shown that it minimizes $J(u,\Delta)$ among all functions u of class $\mathfrak{P}_2$ on Δ which coincide with u_0 on Δ^*. If $g_i=k_i=0$ $(i=1,\cdots,N)$, we may show as in (16.1) that a minimizing function u_0 satisfies (15.3) with $K=M/m$, so that the continuity restrictions obtained for z in §16 follow. If g and k are any functions of class L_2 (whether they satisfy (18.4) or not) the above minimizing argument is valid. If g and k satisfy

$$\iint_{\Delta\cdot C(P,r)}(g^2+k^2)dxdy\leqq Tr^{\pi},\qquad 0<\pi<\mu,\tag{20.1}$$

for all circles with center at a point P in Δ, we conclude as in §16 that a solution u_0 satisfies a condition (15.4) and hence a condition (15.1) for circles with centers at P and therefore obeys the conclusions of Theorem 15.2.

If l_i satisfy (18.4), we may find potential functions

$$V^i(x,y)=\frac{1}{4\pi}\iint_{\Delta}\log\,[(\xi-x)^2+(\eta-y)^2]l_i(\xi,\eta)d\xi d\eta\tag{20.2}$$

which can be shown to be of class $\mathfrak{P}_2'$ and to satisfy a condition (17.3) on any region in the plane and to satisfy

$$\int_{R^*}V_x^i dy-V_y^i dx=\iint_{\Delta\cdot R}l_i dxdy\tag{20.3}$$

for almost all rectangles R in the plane. We then can solve (18.1) with given boundary values with only the d_{ij}, e_{ij}, and $f_{ij}=0$ by solving it with l_i also zero and g_i and k_i replaced by $g_i-V_x^i$ and $k_i-V_y^i$. The solution is easily shown to be unique.

21. We now let G be an arbitrary bounded region and let $\mathfrak{B}_2$ denote the linear subspace of $\mathfrak{P}_2$ consisting of functions u in $\mathfrak{P}_2$ which vanish on G^*. For such functions u, we define the linear operator Tu as the unique solution U in $\mathfrak{B}_2$ of

$$\int_{R^*}(a_{i\beta}U_x^\beta + b_{i\beta}U_y^\beta + d_{i\beta}u^\beta)dy - (b_{\beta i}U_x^\beta + c_{i\beta}U_y^\beta + e_{i\beta}u^\beta)dx = \iint_R (d_{\beta i}u_x^\beta + e_{\beta i}u_y^\beta + f_{i\beta}u^\beta)dxdy.$$

It can be shown that Tu is completely continuous over $\mathfrak{B}_2$ (not $\mathfrak{P}_2$ unless Δ is restricted) and even carries any function u in $\mathfrak{B}_2$ into a function U which is continuous on Δ and can be bounded as in (15.2), the λ there being any number less than $\frac{1}{2}\mu$ (μ in (18.4) and the M depending on the norm of u and of the operator T). Then from the Riesz theory [23] of generalized integral equations, we can conclude that the transformation

$$u - \rho Tu = \phi \tag{21.1}$$

has a linear inverse if ρ is not one of a set of isolated characteristic values. If ρ is not a characteristic value, we can solve the equations (18.1), with the parameter ρ introduced as a multiplier of all the d_{ij}, e_{ij}, or f_{ij}, as follows: First let ϕ be the solution of (18.1) with $\rho=0$ and then solve (21.1) for u in terms of ϕ; the function u so obtained is immediately seen to be the desired solution. In doing this, the g_i and k_i may be merely functions of class L_2 on Δ, but we assume that the d_{ij}, e_{ij}, f_{ij}, and l_i satisfy (18.4). As in §20, if g and k satisfy (20.1) at a point P, we may draw similar conclusions, where the bound for $|\bar{u}(P)|$ given in (15.2) depends only on the quantities m, M, P, μ, T, π, δ, a, the norm of T, and the norm of the inverse operator of (21.1), δ and a having their significance in §15.

22. Let us now suppose that ρ is not a characteristic value and let P_0 be any point in Δ. Let us consider the solution u of (18.1) where $l=0$ and g and k are in L_2 on $\Delta - C(P_0, \sigma)$ and zero in $C(P_0, \sigma)$; such functions satisfy any condition of the type (20.1) for circles $C(P_0, r)$. Clearly the value of $\bar{u}(P_0)$ is a linear functional defined on the space $L_2[\Delta - C(P_0, \sigma)]$ for each $\sigma>0$. We therefore conclude [24] the existence of functions $\Delta_1^{ij}(x_0, y_0; x, y)$ and $\Delta_2^{ij}(x_0, y_0; x, y)$ such that

$$\bar{u}^i(x_0, y_0) = \iint_\Delta (\Delta_1^{i\beta} g_\beta + \Delta_2^{i\beta} k_\beta)dxdy.$$

From the way these Δ_k^{ij} arise, we can conclude that there exist func-

tions $G^{ij}(x_0, y_0; x, y)$ of class $\mathfrak{B}_1'$ over Δ and of class $\mathfrak{B}_2''$ on $\Delta - C(P_0, \sigma)$ for each $\sigma > 0$ such that $G_x^{ij} = \Delta_1^{ij}$, $G_y^{ij} = \Delta_2^{ij}$. It can also be shown that the G^{ij} possess other properties which lead us to call them the *Green's matrix* for (18.1). Moreover if the g and k satisfy (20.1) for every P and r, we conclude that the continuous representative $\bar{u}$ of the solution u of (18.1) in $\mathfrak{B}_2$ is represented everywhere by

$$(22.1)\quad \bar{u}^i(x_0, y_0) = \iint_\Delta (G_x^{i\beta} g_\beta + G_y^{i\beta} k_\beta + G^{i\beta} l_\beta) dx dy, \quad i = 1, \cdots, N,$$

and $\bar{u}^i$ satisfies a uniform Hölder condition on any region H with $\overline{H}$ interior to Δ, this Hölder condition depending only on the distance of $\overline{H}$ from Δ^*, the quantities m, M, P, μ, T, π, a, and any upper bounds for δ and the norms T and the inverse in (21.1).

23. Finally, if (u, v, V) is any solution of (18.6) on a circle $C(P_1, b)$ which satisfies (18.7) there, we can, if ρ is not characteristic, express u in a circle $C(P_1, c)$ with $c < b$ in the form

$$(23.1)\quad \bar{u}^i(x_0, y_0) = \iint_{C(P_1, b) - C(P_1, d)} (G_x^{i\beta} g_\beta + G_y^{i\beta} k_\beta + G^{i\beta} l_\beta) dx dy,$$

$$(x_0, y_0) \in C(P_1, c), \; c < d < b,$$

in which

$$g_i = A_{i\beta}^1 u^\beta + B_{i\beta}^1 v^\beta + C_{i\beta}^1 V^\beta, \qquad k_i = A_{i\beta}^2 u^\beta + B_{i\beta}^2 v^\beta + C_{i\beta}^2 V^\beta,$$

$$l_i = A_{i\beta}^3 u^\beta + B_{i\beta}^3 v^\beta + C_{i\beta}^3 V^\beta$$

where the A_{ij}^k, B_{ij}^k, and C_{ij}^k are bounded and measurable and independent of (x_0, y_0). Also there exists a number $\delta_0 > 0$ which depends only on m, M, P, and μ such that if $\delta \leqq \delta_0$, the norm of T is less than or equal to $\frac{1}{2}$ and the norm of the inverse of (21.1) is less than or equal to 2 for all $\rho \leqq 1$. Thus $\rho = 1$ is surely not a characteristic value if $b \leqq \frac{1}{2}\delta_0$ and the representation (23.1) satisfies a uniform Hölder condition in $C(P_1, c)$ which depends only on m, M, P, μ, Q, ν (of (18.7)), and $b - c$, if $b \leqq \delta_0$. Thus Theorem 19.1 then follows, for, if the solution is given on Δ and H is any subregion with $\overline{H} \subset \Delta$, we may choose b and c so that $\overline{H}$ can be covered by a finite number of circles $C(P_i, c)$ where $C(P_i, b)$ belongs to Δ for each i.

Bibliography

1. The new results stated in this address (and some other results) are to be found in the following eight papers, all but the first of which are by the author. The

first two appear in the Duke Mathematical Journal, vol. 6 (1940), pp. 170–186, 187–215, and the other six will probably appear shortly in the University of California Mathematical Series.

(i) *Functions of several variables and absolute continuity* I, by J. W. Calkin.

(ii) and (iii) the same, *parts* II *and* III.

(iv) *The existence of functions minimizing certain multiple integrals.*

(v) *Lemmas in potential theory.*

(vi) *A general form of Haar's lemma.*

(vii) *Elliptic linear self adjoint systems of partial differential equations in two independent variables.*

(viii) *The differentiability of functions minimizing certain double integrals.*

Numbers (iv) and (viii) have been presented to the Society (February 26, 1938, and April 16, 1938), number (iv) under the title *The existence of solutions of certain minimum problems for multiple integrals.*

2. L. Tonelli, *Sur la semi-continuité des intégrals doubles du calcul des variations*, Acta Mathematica, vol. 53 (1929), pp. 325–346; and *L'estremo assoluto degli integrali doppi*, Annali della Scuola Normale Superiore di Pisa, (2), vol. 2 (1933), pp. 89–130.

3. E. J. McShane, *Integrals over surfaces in parametric form*, Annals of Mathematics, (2), vol. 34 (1933), pp. 815–838; and *Existence theorems for double integral problems of the calculus of variations*, Transactions of this Society, vol. 38 (1935), pp. 549–563.

4. E. Hopf, *Zum analytischen Charakter der Lösungen regulärer zweidimensionaler Variationsprobleme*, Mathematische Zeitschrift, vol. 30, pp. 404–413.

5. C. B. Morrey, Jr., *On the solutions of quasi-linear elliptic partial differential equations*, Transactions of this Society, vol. 43 (1938), pp. 126–166.

6. G. C. Evans, *Fundamental points of potential theory*, Rice Institute Pamphlets, vol. 7, no. 4 (1920), pp. 252–329, particularly pp. 274, 275. We shall refer to this paper by the letter P.

7. G. C. Evans, *Complements of potential theory, part* II, American Journal of Mathematics, vol. 55 (1933), pp. 29–49, especially pp. 42–46.

8. L. Tonelli, *Sulla quadratura delle superficie*, Atti della Reale Accademia dei Lincei, (6), vol. 3 (1926), pp. 633–638.

9. See, for instance, de la Vallée Poussin, *Intégrales de Lebesgue, Fonctions d'Ensemble, Classes de Baire*, Paris, 1916, pp. 59–61.

10. S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, chap. 4. We shall refer to this book as "Banach."

11. H. Rademacher, *Über partielle und totale Differenzierbarkeit von Funktionen mehrerer Variabeln, und über die Transformation der Doppelintegrale*, Mathematische Annalen, vol. 79 (1918), pp. 340–359.

12. See P [6], p. 282.

13. S. Saks, *On the surfaces without tangent planes*, Annals of Mathematics, (2), vol. 34 (1933), pp. 114–124.

14. R. Courant and D. Hilbert, *Methoden der mathematischen Physik*, Berlin, 1937, vol. 2, chap. 7, §1.

15. Banach, chap 9.

16. Banach, p. 23.

17. R. Courant, *Plateau's problem and Dirichlet's principle*, Annals of Mathematics, (2), vol. 38 (1937), pp. 679–724.

18. A. Haar, *Über die Variation der Doppelintegrale*, Journal für die reine und angewandte Mathematik, vol. 149 (1919), pp. 1–18; and *Über das Plateausche Prob-*

lem, Mathematische Annalen, vol. 97 (1926–1927), pp. 124–158, particularly pp. 146–151.

19. L. Lichtenstein, *Über den analytischen Charakter der Lösungen regulärer zweidimensionaler Variationsprobleme*, Bulletin de l'Académie des Sciences de Cracovie, Classe des Sciences Mathématiques et Naturelles, (A), December, 1912, pp. 915–941.

20. S. Bochner, *Harmonic surfaces in Riemann metric*, this Bulletin, vol. 45 (1939), p. 832.

21. J. Leray and J. Schauder, *Topologie et équations fonctionelles*, Annales Scientifiques de L'École Normale Supérieure, (3), vol. 51 (1934), pp. 45–78.

22. C. B. Morrey, Jr., *The problem of Plateau on a Riemannian manifold*, this Bulletin, abstract 45-9-350.

23. Banach, chap. 10. For the original see F. Riesz, *Über lineare Funktionalgleichungen*, Acta Mathematica, vol. 41 (1916), pp. 71–98.

24. Banach, chap. 4.

UNIVERSITY OF CALIFORNIA

NULLIFYING FUNCTIONS[1]

F. J. MURRAY

Introduction. A function $f(x)$ defined on the unit interval (0, 1) will be called nullifying if we can find a set S of (0, 1) for which $m(S)=1$, $m(\{f(x); x \in S\})=0$. Examples of homeomorphisms which are nullifying and hence termed singular are well known.[2] We shall however consider simply the nullifying property itself.

If $f(x)$ is nullifying and $\phi(x)$ is not, one might expect that $\phi(x)+f(x)$ shares with ϕ the property of being not nullifying. But this is not always true as the following example shows. Let $.\alpha_1\alpha_2\cdots$ denote the dyadic expansion for x, that is, $x=\alpha_1/2+\alpha_2/2^2+\cdots$ with $\alpha_i=0$ or 1. Let $v_1(x)=.\alpha_1 0\alpha_3 0\cdots$ and $v_2(x)=.0\alpha_2 0\alpha_4\cdots$. It is easily verified that both v_1 and v_2 are nullifying. Hence $f(x)=1-v_1$ is also nullifying. Let $\phi=x$. Then $f+\phi=1-v_1+x=1+v_2$ is also nullifying.

But this suggests the question: Does there exist a nullifying function $f(x)$ such that $f(x)+\rho x$ is nullifying for every value of ρ? We construct such a function in the present note.

Our method of proof can be summarized as follows. Considering the set $\{f(x)+\rho x; x \in (0, 1)\}$, we let $\rho=\cot\theta$. (Note $\theta\neq 0$.) If this set has measure zero, this will still be true if we multiply by $\sin\theta$ and conversely. Thus we may consider the sets $\{x\cos\theta+f(x)\sin\theta; x \in (0, 1)\}$ for each $\theta\neq 0$ between $-\pi/2$ and $\pi/2$. If we consider the line through the origin of inclination θ, we can assign a coordinate to each of its points in the usual manner with positive direction to the right or, in the case of the y axis, upwards. The set $\{x\cos\theta+f(x)\sin\theta; x \in (0, 1)\}$ is the set of coordinates of the projection onto this line of the graph of $f(x)$. Thus it suffices to find a function $f(x)$ which is such that the projection of its graph onto any line not parallel to the x axis is of measure zero. We proceed to find the graph of such a function by an intersection process on sets in the plane. This process is described in detail in what follows.

A more general question is: Given $F(x, y, \rho)$, under what circumstances can we find a function $f(x)$ such that $F(x, f(x), \rho)$ is nullifying in x for every value of ρ? It is comparatively easy to abstract the properties of $F=y+\rho x$ which are essential to the present discussion, and these will prove sufficient to obtain an answer to the question.

[1] Presented to the Society, December 29, 1939.

[2] Cf. E. R. van Kampen and Aurel Wintner, *On a singular monotone function*, Journal of the London Mathematical Society, vol. 12 (1937), pp. 243–244. References to preceding examples are given in this paper.

However the writer hopes to obtain a more penetrating analysis of this subject soon. The writer will also consider the more general question of obtaining an $f(x)$ the substitution of which will make nullifying not only one but a set of F's.

Definition of a G_s. Let us divide evenly the unit square of the plane in squares of side $1/2^s$. The coordinates of the vertices of these squares are dyadic rational, that is, of the form $a/2^s$ for some integer a. Two squares are said to be in the same column if they have the same projection on the x axis. A set of such squares with one and only one (closed) square in each column will be termed[3] a G_s.

Notation. If l_θ is the line through the origin with inclination θ and P is any point of the plane, we shall denote by $\pi_\theta(P)$ the projection of P onto l_θ. If S is a set of points, we denote by $\pi_\theta(S)$ the projection of S onto l_θ. The linear measure of a linear set S we denote by $\mu(S)$.

LEMMA 1. *Let m be such that $0 \leqq m \leqq 1$. Let $\epsilon > 0$ and a G_s be given. Let α be such that $\pi/2 \geqq \alpha \geqq \pi/4$ and $\cot \alpha = m$. Then we can find a $G_{s+t} \subset G_s$ such that for $0 \leqq \theta \leqq \pi/2$,*

$$\mu(\pi_\theta(G_{s+t})) \leqq (1 + m^2)^{1/2} \sin \left| \theta - \alpha \right| + \epsilon.$$

PROOF. Let us consider a square B of G_s. Then it is possible to take a line $l^{(1)}$ of slope $-m$, such that for every x in the projection of B on the x axis, we have a point (x, y) in B and on $l^{(1)}$. Let $l^{(1)}$ intersect the left-hand side of B at P_B and the right-hand side at Q_B.

Let us divide B into squares of side $1/2^{s+t}$. Now it is readily seen that for each column of smaller squares in B there is at least one square with an interior point on $l^{(1)}$. (If $m = 0$, this is not true, but then we can substitute for "interior point," "interior point or point on the upper side.") Define G_{s+t} so as to contain for each column the lowest such square.

We also take a square of side $1/2^{s+t}$, whose upper right-hand vertex is at P_B. We denote the lower left-hand end point of this square by P_B'. Similarly we take a square of side $1/2^{s+t}$ whose lower left-hand vertex is Q_B, and we denote the upper right-hand vertex by Q_B'.

We are going to consider $\pi_\theta(G_{s+t})$ and we let $<$ and $\leqq$ refer to the order of the points on l_θ, the direction of the greater being to the right of the smaller.

We shall show that if $P \in G_{s+t}$, and $0 \leqq \theta \leqq \alpha$, then $\pi_\theta(P_B') \leqq \pi_\theta(P)$.

[3] We shall ignore the logical distinction between a set of squares and the corresponding set of points. G_s denotes either a (closed) set of points or a set of squares according to the context.

For P_B' (x_B', y_B') is the lower left-hand vertex of a square whose upper right-hand vertex alone is on $l^{(1)}$. Now every square C of G_{s+t} contains an interior point which is on $l^{(1)}$ (if $m=0$, confer above). Thus if $l^{(1)\prime}$ is parallel to $l^{(1)}$ and through P_B', every point P (x, y) of G_{s+t} is above or on $l^{(1)\prime}$. Furthermore $x>x_B'$.

Let P' (x', y') be the intersection of $l^{(1)\prime}$ and the line through P parallel to the y axis. P is above P' and hence $\pi_\theta(P)\geqq\pi_\theta(P')$. On the other hand $x'=x>x_B'$. Let $\bar{l}$ denote the line through P_B' perpendicular to l_θ. The slope of $\bar{l}$ is more strongly negative than that of $l^{(1)\prime}$. It follows that for $x\geqq x_B'$, $l^{(1)\prime}$ is to the right of $\bar{l}$ and hence $\pi_\theta(P')\geqq\pi_\theta(P_B')$. This and the previous result yield $\pi_\theta(P)\geqq\pi_\theta(P_B')$.

A similar argument will show that $\pi_\theta(Q_B')\geqq\pi_\theta(P)$.

Thus $\pi_\theta(P_B')\leqq\pi_\theta(P)\leqq\pi_\theta(Q_B')$. This implies that the projection of that part of G_{s+1} which lies in B is contained in the interval $\pi_\theta(P_B')\pi_\theta(Q_B')$. Now if P_B has the coordinates (x_B, y_B), then Q_B has the coordinates $(x_B+1/2^s,\ y_B-m/2^s)$, P_B' has the coordinates $(x_B-1/2^{s+t},\ y_B-1/2^{s+t})$, and Q_B' has the coordinates $(x_B+1/2^s+1/2^{s+t},\ y_B-m/2^s+1/2^{s+t})$.

Consider $l_{\theta+\pi/2}$, the line through the origin perpendicular to l_θ. According to a formula of elementary analytic geometry, the directed distance of any point P (x, y) to $l_{\theta+\pi/2}$ is $x\cos\theta+y\sin\theta$, the directed distance being positive if (x, y) is to the right of $l_{\theta+\pi/2}$ and negative to the left. This directed distance is not changed if we project P onto l_θ, and so $x\cos\theta+y\sin\theta$ represents also the directed distance of $\pi_\theta(P)$ from $l_{\theta+\pi/2}$, or since l_θ is perpendicular to $l_{\theta+\pi/2}$, from the origin O along l_θ.

Thus $O\pi_\theta(P_B')$ has length $(x_B-1/2^{s+t})\cos\theta+(y_B-1/2^{s+t})\sin\theta$ and $O\pi_\theta(Q_B')$ has the length $(x_B+1/2^s+1/2^{s+t})\cos\theta+(y_B-m/2^s+1/2^{s+t})\sin\theta$. Hence $\pi_\theta(P_B')\pi_\theta(Q_B')$ has length

$$(1/2^s+1/2^{s+t-1})\cos\theta-(m/2^s-1/2^{s+t-1})\sin\theta$$

for $0\leqq\theta\leqq\alpha$. Thus the projection of that part of G_{s+1} which lies in B has measure not greater than

$$1/2^s(\cos\theta-m\sin\theta)+1/2^s(1/2^{t-1}(\cos\theta+\sin\theta)$$
$$<1/2^s(\cos\theta-m\sin\theta)+1/2^s(1/2^{t-2}).$$

There are 2^s squares like B, and this implies that for $0<\theta\leqq\alpha$, the projection of G_{s+t} has measure not greater than

$$(\cos\theta-m\sin\theta)+1/2^{t-2}=(1+m^2)^{1/2}\sin(\alpha-\theta)+1/2^{t-2}.$$

A similar argument holds in the case $\alpha<\theta\leqq\pi/2$, with however the smaller added squares in different positions, and during the argument

certain inequalities are reversed. The corresponding result is that, for $\alpha<\theta\leqq\pi/2$,

$$\mu(\pi_\theta(G_{s+t})) < m\sin\theta-\cos\theta+1/2^{t-2}$$
$$=(1+m^2)^{1/2}\sin(\theta-\alpha)+1/2^{t-2}.$$

If t is taken large enough so that $1/2^{t-2}<\epsilon$, the final results of the preceding two paragraphs are sufficient to prove our lemma.

LEMMA 2. *Suppose $m>1$, and that a G_s and an $\epsilon>0$ are given. Let α be such that $\pi/4>\alpha>0$, $\cot\alpha=m$. Then we can find a $G_{s+t}\subset G_s$ such that for $0\leqq\theta<\pi/2$,*

$$\mu(\pi_\theta(G_{s+t}))\leqq(1+m^2)^{1/2}\sin|\theta-\alpha|+\epsilon.$$

PROOF. Let B be any square of G_s. Let $l^{(1)}, l^{(2)}, \cdots, l^{(k)}$ denote a set of lines of slope $-m$ with the following properties: (a) each $l^{(i)}$ contains an interior point of B; (b) if $P_B^{(i)}Q_B^{(i)}$ is the line segment $B\cdot l^{(i)}$ (B is closed), $P_B^{(1)}$ is on the left side of B, $Q_B^{(k)}$ is on the right side of B, and for $i=1, 2, \cdots, k-1$, $P_B^{(i+1)}$ is on the upper side of B, $Q_B^{(i)}$ is on the lower side of B, and $Q_B^{(i)}$ has the same x coordinate as $P_B^{(i+1)}$. The existence of such a set of lines is easily shown.

Let the coordinates of $P_B^{(i)}$ be $(x_B^{(i)}, y_B^{(i)})$, those of $Q_B^{(i)}$ be $(u_B^{(i)}, v_B^{(i)})$. From the above we see that $x_B^{(1)}=a/2^s$, $y_B^{(1)}=b/2^s+\rho/2^s$ where $0<\rho\leqq1$, $u_B^{(1)}=a/2^s+\rho/m\cdot2^s$, $u_B^{(2)}=a/2^s+(\rho+1)/m\cdot2^s$, and in general $u_B^{(i)}=a/2^s+(\rho+i-1)/2^s\cdot m$, for $i=1, 2, \cdots, k-1$, while $u_B^{(k)}=(a+1)/2^s$. From this it follows that

$$(\rho+k-2)/m\cdot2^s<1/2^s\leqq(\rho+k-1)/m\cdot2^s$$

or $(\rho+k-2)<m<\rho+k-1$ or $k-1<m+(1-\rho)\leqq k$. Hence if $[m]$ denotes the largest integer less than or equal to m, then $k=[m]$, $[m]+1$, or $[m]+2$. Thus $k\leqq m+2$.

We next divide the square B of S_s into smaller squares of side $1/2^{s+t}$. The lines $l^{(1)}, \cdots, l^{(k)}$ have been chosen so that for each x in the projection of B on the x axis, there is at least one $l^{(i)}$ which contains a point (x, y) in B. This can be used to show that for each column of smaller squares in B, there is at least one smaller square which contains a point on some $l^{(i)}$ and interior to B. Define G_{s+t} so as to contain the lowest such square in each column.

For each i, let us consider those squares of G_{s+t} which contain points of $l^{(i)}$. An argument similar to that used in Lemma 1 will show that the projection of this part of G_{s+t} on l_θ has measure not greater than

$$(\text{length of } l^{(i)}\cdot B)\sin|\theta-\alpha|+1/2^{s+t-2}.$$

The sum of the lengths of the $l^{(i)} \cdot B$ is $(1/2^s)(1+m^2)^{1/2}$ and there are at most $[m]+2$ of them. Thus

$$\mu(\pi_\theta(G_{s+t} \cdot B)) \leqq (1/2^s)((1+m^2)^{1/2} \sin |\theta - \alpha| + ([m]+2)/2^{t-2}).$$

Since there are 2^s such squares B, we have

$$\mu(\pi_\theta(G_{s+t})) \leqq (1+m^2)^{1/2} \sin |\theta - \alpha| + ([m]+2)/2^{t-2}.$$

We can take t sufficiently large to obtain our result.

This result is interesting only in the range $0<\theta<2\alpha$, for outside this range other methods give more effective inequalities.

The argument of Lemmas 1 and 2 can be modified to apply to the case in which m is negative, with the following result.

LEMMA 3. *Suppose an $m<0$, a G_s and an $\epsilon>0$ are given. Let α be such that $\cot \alpha = m$ and $-\pi/2 \leqq \alpha < 0$. Then we can find a $G_{s+t} \subset G_s$ such that for $-\pi/2 \leqq \theta < 0$,*

$$\mu(\pi_\theta(G_{s+t})) \leqq (1+m^2)^{1/2} \sin |\theta - \alpha| + \epsilon.$$

LEMMA 4. *Let m_1 and m_2 be given with $m_2 > m_1 \geqq 0$. Let α_i be such that $\pi/2 \geqq \alpha_i > 0$, $\cot \alpha_i = m_i$ for $i=1, 2$. Consider the function*

$$F(\theta) = \min ((1+m_2^2)^{1/2} \sin (\theta - \alpha_2), (1+m_1^2)^{1/2} \sin (\alpha_1 - \theta))$$

for $\alpha_2 \leqq \theta \leqq \alpha_1$. Then for this range of θ,

$$F(\theta) \leqq (m_2 - m_1)/2(m_1^2+1)^{1/2}.$$

PROOF. For $\alpha_2 \leqq \theta \leqq \alpha_1$, $(1+m_2^2)^{1/2} \sin (\theta - \alpha_2)$ is increasing while $(1+m_1)^{1/2} \sin (\alpha_1 - \theta)$ is decreasing. Also we readily see that if η is such that

$$(1+m_2^2)^{1/2} \sin (\eta - \alpha_2) = (1+m_1^2)^{1/2} \sin (\alpha_1 - \eta)$$

then $F(\theta)$ is equal to the first expression for $\alpha_2 \leqq \theta \leqq \eta$ and to the second expression for $\eta \leqq \theta \leqq \alpha_1$. Hence $F(\eta)$ is the maximum value of $F(\theta)$ for the given range of θ.

Expanding $\sin (\eta - \alpha_2)$ and $\sin (\alpha_1 - \eta)$, using the definition of α_i, collecting terms in $\sin \eta$ and $\cos \eta$ and dividing by $\cos \eta$ yields

$$\tan \eta = 2/(m_1 + m_2).$$

The value of $F(\eta)$ is then seen to be

$$(m_2 - m_1)/2(1 + (m_1+m_2)^2/4)^{1/2} \leqq (m_2 - m_1)/2(1+m_2^2)^{1/2}.$$

This and the result of the preceding paragraph prove the lemma.

Consider the sequence $0, 1, 0, -1, 2, 3/2, 1, 1/2, 0, -1/2, -1, -3/2, -2, 4, 15/4, \cdots, 1/4, 0, -1/4, \cdots, -15/4, -4, \cdots$. We

denote the ith term in this sequence by m_i. Notice that the terms of this sequence can be grouped so that the first group contains 1, the second group 3, the third group 9, and the kth group $2^{2k-3}+1$. The maximum value in the kth group is 2^{k-2} and the difference between any two adjacent terms in this group is $1/2^{k-2}$.

Let $\{\epsilon_i\}$ denote the sequence such that if i is a subscript of the kth group of the preceding paragraph, then $\epsilon_i=1/2^k$. Let $G^{(0)}=G_0$, the unit square itself, and $G^{(i)}$ be defined as the G_{s+t}, which results when either Lemma 1, 2, or 3 (depending on m_i) is applied to $G^{(i-1)}$, m_i and ϵ_i.

LEMMA 5. *Let $G^{(i)}$ be as above. Let θ be such that $\pi/2 \geqq \theta > 0$. Then* $\lim_{i\to\infty} \mu(\pi_\theta(G^{(i)}))=0$.

PROOF. Let $\epsilon>0$ be given. Take k such that $\epsilon/2>1/2^{k-1}$ and $2^{k-2}>\cot\theta$. Then we can find an m_i in the kth group such that $m_i>0$ and $m_i \geqq \cot\theta > m_{i+1}$.

Now since $G^{(i)} \supset G^{(i+1)}$, we have $\pi_\theta(G^{(i)}) \supset \pi_\theta(G^{(i+1)})$ and $\mu(\pi_\theta(G^{(i)})) \geqq \mu(\pi_\theta(G^{(i+1)}))$. Thus $\mu(\pi_\theta(G^{(i+1)}))$ is subject to the inequality which Lemmas 1 and 2 impose upon $\mu(\pi_\theta(G^{(i)}))$. Thus

$$\mu(\pi_\theta(G^{(i+1)})) \leqq \min\left[(1+m_i^2)^{1/2}\sin\left|\theta-\alpha_i\right|, (1+m_{i+1}^2)^{1/2} \cdot \sin\left|\theta-\alpha_{i+1}\right|\right]+\epsilon_i, \qquad (\cot\alpha_i=m_i).$$

Since $\alpha_i \leqq \theta \leqq \alpha_{i+1}$, Lemma 4 yields

$$\mu(\pi_\theta(G^{(i+1)})) \leqq (m_i-m_{i+1})/2(1+m_{i+1}^2)^{1/2}+\epsilon_i < 1/2^{k-1}+1/2^k < \epsilon.$$

Also if $j \geqq i+1$, $G^{(j)} \subset G^{(i+1)}$, and we obtain

$$0 < \mu(\pi_\theta(G^{(j)})) \leqq \mu(\pi_\theta(G^{(i+1)})) < \epsilon.$$

Thus we have shown that given an $\epsilon>0$, we can find an i such that for $j>i$ this equation holds. The lemma is now proved.

LEMMA 6. *Let $G^{(i)}$ be as in Lemma 5. Let $G=\prod G^{(i)}$. Then G is a non-empty closed set such that if $0<\theta\leqq\pi/2$, then $\mu(\pi_\theta(G))=0$.*

PROOF. Since the $G^{(i)}$'s form a decreasing sequence of closed sets, their intersection is a non-empty closed set. Since $G \subset G^{(i)}$, $0 \leqq \mu(\pi_\theta(G)) \leqq \mu(\pi_\theta(G^{(i)}))$. Thus Lemma 5 now implies $\mu(\pi_\theta(G))=0$.

Results similar to those of Lemmas 5 and 6 hold for $0>\theta>-\pi/2$. The method of obtaining them should be clear from the preceding discussion and we merely state the final result as follows.

LEMMA 7. *If G is as in Lemma 6, then for $0>\theta>-\pi/2$, $\mu(\pi_\theta(G))=0$.*

LEMMA 8. *G is the graph of a function $y=f(x)$.*

PROOF. For $0 \leqq a \leqq 1$, let p_a denote the line $x=a$. Then the sets $p_a \cdot G^{(i)}$ form a decreasing sequence of closed intervals on p_a, with one and only one point (a, b) in common. Thus $p_a \cdot G$ consists of one and only one point (a, b) and b is a function of a.

Now, as we have pointed out in the proof of Lemma 1, for any point $P(x, y)$, $x \cos\theta + y \sin\theta$ is the directed distance of $\pi_\theta(P)$ along l_θ from the origin. Using Lemmas 6, 7, and 8, we obtain that for $\pi/2 \geqq \theta > 0$ or $0 > \theta > -\pi/2$,

$$\begin{aligned} 0 &= \mu(\pi_\theta(G)) = \mu(\{x \cos\theta + y \sin\theta;\ (x, y) \in G\}) \\ &= \mu(\{x \cos\theta + f(x) \sin\theta;\ 0 \leqq x \leqq 1\}) \\ &= |\sin\theta|\, \mu(\{x \cot\theta + f(x);\ 0 \leqq x \leqq 1\}). \end{aligned}$$

Letting $\rho = \cot\theta$ we obtain that for every value of ρ,

$$\mu(\{f(x) + \rho x;\ 0 \leqq x \leqq 1\}) = 0.$$

Since $f(x)$ is the limit of step functions, it is measurable, and we have proved the following theorem.

THEOREM. *There exists a measurable function $f(x)$ defined for $0 \leqq x \leqq 1$, such that for every value of ρ, $f(x) + \rho x$ is nullifying.*

COLUMBIA UNIVERSITY

CERTAIN SELF-RECIPROCAL FUNCTIONS

BRIJ MOHAN

In 1932 and 1933 I [5, 6] gave some rules connecting different classes of self-reciprocal functions. The object of this note is to derive some new self-reciprocal functions with the help of those rules.

I will say that a function is R_ν if it is self-reciprocal for J_ν transforms, where $\nu > -1$.

I will make use of the following results given in the papers referred to:

If $f(x)$ is R_μ, the functions $g(x)$ given by the following integral formulas are all R_ν:

(i) $$g(x) = x^{(\nu-\mu+1)/2}\int_0^\infty y^{(\nu-\mu+1)/2}J_{(\mu+\nu)/2}(xy)f(y)dy,$$

(ii) $$g(x) = x^{(\mu-\nu+1)/2}\int_0^\infty y^{(\mu-\nu+1)/2}J_{(\mu+\nu)/2}(xy)f(y)dy,$$

(iii) $$g(x) = \int_0^\infty \frac{y^{\mu+1/2}f(xy)}{(1+y^2)^{1+\mu/2+\nu/2}}\,dy,$$

(iv) $$g(x) = \int_1^\infty \frac{y^{1/2-\mu}f(xy)}{(y^2-1)^{1-\mu/2+\nu/2}}\,dy,$$

(v) $$g(x) = \int_0^1 \frac{y^{1/2+\mu}f(xy)}{(1-y^2)^{1+\mu/2-\nu/2}}\,dy.$$

If, in (ii) we take the familiar R_μ function

(1) $$x^{\mu+1/2}e^{-x^2/2}$$

for $f(x)$, we get

$$g(x) = x^{(\mu-\nu+1)/2}\int_0^\infty y^{(\mu-\nu+1)/2}J_{(\mu+\nu)/2}(xy)\cdot y^{\mu+1/2}e^{-y^2/2}dy$$

$$= x^{(\mu-\nu+1)/2}\int_0^\infty y^{3\mu/2-\nu/2+1}e^{-y^2/2}J_{\mu/2+\nu/2}(xy)dy.$$

Evaluating this integral by Hankel's formula [7], we get

$$g(x) = x^{(\mu-\nu+1)/2}\frac{\Gamma(\mu+1)(x/2^{1/2})^{\mu/2+\nu/2}}{2^{-3\mu/4+\nu/4}\Gamma(1+\mu/2+\nu/2)}e^{-x^2/2}$$
$$\cdot {}_1F_1(\nu/2-\mu/2;\,1+\mu/2+\nu/2;\,x^2/2), \qquad \mu > -1.$$

This shows that the function $x^{\mu+1/2}e^{-x^2/2}\,{}_1F_1(\nu/2-\mu/2;1+\mu/2+\nu/2;x^2/2)$, $(\mu>-1)$, which is the same as

$$x^{\nu+2n+1/2}e^{-x^2/2}\,{}_1F_1(-n;n+\nu+1;x^2/2), \qquad \nu+2n>-1, \tag{2}$$

is R_ν.

Particular cases. (α) Since, when n is a positive integer,

$$L_n^{(\alpha)}(x)=\frac{\Gamma(1+\alpha+n)}{n!\Gamma(\alpha+1)}\,{}_1F_1(-n;\alpha+1;x),$$

where $L_n^{(\alpha)}(x)$ denotes the generalized Laguerre polynomial of order n, it follows that the function $x^{\nu+2n+1/2}e^{-x^2/2}L_n^{(n+\nu)}(x^2/2)$, $(\nu+2n>-1)$, is R_ν for positive integral values of n.

This result was given by Howell [4].

(β) Since, when n is a positive integer,

$$T_\mu^n(x)=\frac{(-1)^n}{n!\Gamma(1+\mu)}\,{}_1F_1(-n;1+\mu;x),$$

where $T_\mu^n(x)$ is Sonine's polynomial of order n, it follows that the function

$$x^{\nu+2n+1/2}e^{-x^2/2}T_{n+\nu}^n(x^2/2), \qquad \nu+2n>-1,$$

is R_ν for positive integral values of n.

(γ) If we put $n=-\nu-1/2$, our function becomes $x^{-\nu-1/2}e^{-x^2/2}\cdot{}_1F_1(\nu+1/2;1/2;x^2/2)$.

As a particular case, when ν is an integer, this becomes a constant multiple of

$$x^{-\nu-1/2}e^{-x^2/4}D_{-2\nu-1}(x). \tag{3}$$

(δ) When $n=1/2-\nu$, the function is $x^{-\nu+3/2}e^{-x^2/2}\,{}_1F_1(\nu-1/2;3/2;x^2/2)$, $(\nu<2)$.

If ν is an integer, this is a constant multiple of

$$x^{-\nu+1/2}e^{-x^2/4}D_{-2\nu+1}(x). \tag{4}$$

The functions (3) and (4) were given by Mitra [2].

Taking the same function (1) for $f(x)$, we have, from (iii),

$$g(x)=\int_0^\infty \frac{y^{\mu+1/2}e^{-x^2y^2/2}(xy)^{\mu+1/2}}{(1+y^2)^{1+\mu/2+\nu/2}}\,dy = x^{\mu+1/2}\int_0^\infty \frac{y^{2\mu+1}e^{-x^2y^2/2}}{(1+y^2)^{1+\mu/2+\nu/2}}\,dy.$$

Evaluating this integral by a formula given by Whittaker [8] we find that $g(x)$ is a constant multiple of

$$x^{(\mu+\nu-1)/2}e^{x^2/4}W_{-1/2-\mu/4-\nu/4,(\nu-\mu)/4}(x^2/2),$$

where $\mu > -1$.

Thus, we see that the function

$$x^{\nu-2m-1/2}e^{x^2/4}W_{2m-\nu-1/2,m}(x^2/2), \tag{5}$$

where $\nu-4m > -1$, is R_ν.

This function was given by Bailey [1].[1]

If we put the same function (1) for $f(x)$ in (i) and evaluate the corresponding integral by Hankel's formula [7], we arrive back at the R_ν function

$$x^{\nu+1/2}e^{-x^2/2}. \tag{6}$$

If, in (iv), we take the same function for $f(x)$ we arrive at the function (6) again.

If we take the same function for $f(x)$ in (v), we again arrive at the function (2).

References

1. W. N. Bailey, *Some classes of functions which are their own reciprocals in the Fourier-Bessel integral transform*, Journal of the London Mathematical Society, vol. 5 (1930), pp. 258–265, (6.3).

2. ———, *Self-reciprocal functions involving confluent hyper-geometric functions*, ibid., vol. 13 (1938), pp. 111–112.

3. S. C. Dhar, *On certain functions which are self-reciprocal in the Hankel transform*, ibid., vol. 14 (1939), pp. 30–32.

4. W. T. Howell, *A note on Laguerre polynomials*, Philosophical Magazine, (7), vol. 23 (1937), pp. 807–811.

5. Brij Mohan (formerly B. M. Mehrotra), *Some theorems on self-reciprocal functions*, Proceedings of the London Mathematical Society, (2), vol. 34 (1932), pp. 231–240, §8.

6. ———, *On some self-reciprocal functions*, Bulletin of the Calcutta Mathematical Society, vol. 25 (1933), pp. 167–172, §2.

7. G. N. Watson, *Theory of Bessel Functions*, Cambridge, 1922, §§13.3, (3) and 13.3, (4).

8. E. T. Whittaker and G. N. Watson, *Modern Analysis*, Cambridge, 4th edition, 1927, §16.12.

Benares Hindu University
Benares, India

[1] Recently, Dhar [3] has given another proof of the fact that this function is R_ν. The other R_ν function $x^{\nu+2m-1/2}e^{x^2/4}W_{-2m-\nu-1/2,m}(x^2/2)$ that he gives, is the same as (5) with "$-m$" written for "m."

A REMARK ON THE SUM AND THE INTERSECTION OF TWO NORMAL IDEALS IN AN ALGEBRA

TADASI NAKAYAMA

Let F be a quotient field of a commutative domain of integrity o in which the usual arithmetic holds.[1] Consider an algebra $\mathfrak{A}$ with a unit element over F. Let $\mathfrak{J}_1$, $\mathfrak{J}_2$, $\mathfrak{J}_3$, $\mathfrak{J}_4$ be four arbitrary maximal orders in $\mathfrak{A}$ and $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ be three arbitrary normal ideals. We prove the following theorems.

THEOREM 1. *If* $\mathfrak{J}_1 \cap \mathfrak{J}_2 = \mathfrak{J}_3 \cap \mathfrak{J}_4$ [*or* $(\mathfrak{J}_1, \mathfrak{J}_2) = (\mathfrak{J}_3, \mathfrak{J}_4)$], *then either* $\mathfrak{J}_1 = \mathfrak{J}_3$, $\mathfrak{J}_2 = \mathfrak{J}_4$ *or* $\mathfrak{J}_1 = \mathfrak{J}_4$, $\mathfrak{J}_2 = \mathfrak{J}_3$.

THEOREM 2. *Both the left and the right orders of* $(\mathfrak{J}_1, \mathfrak{J}_2)$ *are* $\mathfrak{J}_1 \cap \mathfrak{J}_2$. *Also* $\mathfrak{J}_1 \cap \mathfrak{J}_2 \subseteqq \mathfrak{J}_3$ *if and only if* $(\mathfrak{J}_1, \mathfrak{J}_2) \supseteqq \mathfrak{J}_3$; *if this is the case the distance ideal* $\mathfrak{d}_{21}$ *of* $\mathfrak{J}_2$ *to* $\mathfrak{J}_1$ *is divisible by the distance ideal*[2] $\mathfrak{d}_{31}$ *of* $\mathfrak{J}_3$ *to* $\mathfrak{J}_1$.

THEOREM 3. *The left, say, order* $\mathfrak{o}$ *of the intersection* $\mathfrak{a} \cap \mathfrak{b}$ [*the sum* $(\mathfrak{a}, \mathfrak{b})$] *is an intersection of two suitable maximal orders.*

More precisely, if $\mathfrak{r}$ and $\mathfrak{s}$ are normal ideals such that $\mathfrak{b} = \mathfrak{r}\mathfrak{a}\mathfrak{s}$ in the sense of proper multiplication and if $\mathfrak{t}$ is the smallest two-sided ideal of the right order of $\mathfrak{a}$ which divides $\mathfrak{s}$ while $\mathfrak{t}'$ is the largest two-sided ideal of the same maximal order which is divisible by $\mathfrak{s}$, then $\mathfrak{o}$ is the intersection of the left orders of the two normal ideals $\mathfrak{a} \cap \mathfrak{r}\mathfrak{a}\mathfrak{t}$ and $\mathfrak{a} \cap \mathfrak{r}\mathfrak{a}\mathfrak{t}'$ [$(\mathfrak{a}, \mathfrak{r}\mathfrak{a}\mathfrak{t})$ and $(\mathfrak{a}, \mathfrak{r}\mathfrak{a}\mathfrak{t}')$].[3] *The left order of* $\mathfrak{a} \cap \mathfrak{b}$ *coincides with the right order of* $(\mathfrak{a}^{-1}, \mathfrak{b}^{-1})$.

THEOREM 4. $\mathfrak{a} \cap \mathfrak{b} \subseteqq \mathfrak{c}$ *implies* $(\mathfrak{a}^{-1}, \mathfrak{b}^{-1}) \supseteqq \mathfrak{c}^{-1}$ *and conversely.*

For the proof we have, according to the well known reduction, only to treat the case where F is a p-adic field $F = F_p$ and $\mathfrak{A}$ is a normal simple algebra over F. Then $\mathfrak{A}$ is a (complete) matric ring $D_r = \sum_{i,k=1}^{r} \epsilon_{ik} D$ over a division algebra D, where ϵ_{ik} is a system of matric units commutative with every element of D. D possesses a unique maximal order I, and I has a unique prime ideal P.

Notation. If a_{ik}, $(i, k = 1, 2, \cdots, r)$, is a system of rational integers, we denote by $M(a_{ik})$ the ideal $\sum_{i,k} \epsilon_{ik} P^{a_{ik}}$ in $\mathfrak{A}$.

---

[1] In the following we shall adopt the terminologies used in M. Deuring, *Algebren*, Ergebnisse der Mathematik, vol. 4, no. 1, 1935.

[2] If the algebra is a quaternion algebra, then the converse is also valid. Cf. M. Eichler, Journal für die reine und angewandte Mathematik, vol. 174 (1936), §7.

[3] Thus the intersection and the sum are no more normal ideals except for trivial cases; cf. Nakayama, Proceedings of the Imperial Academy of Japan, vol. 12 (1936).

$M(a_{ik})$ is an order if and only if $a_{ii}=0$, $a_{il}+a_{lk}\geqq a_{ik}$ for all i, k, l. On assuming this condition it is a maximal order if and only if $\sum a_{ik}=0$. By a simple calculation we then have the following lemma.[4]

LEMMA 1. *A necessary and sufficient condition that $M(a_{ik})$ be a maximal order is that there should exist r rational integers c_i such that $a_{ik}=c_k-c_i$. Every normal ideal whose left and right orders are $M(c_k-c_i)$ and $M(d_k-d_i)$ respectively has the form $P^aM(d_k-c_i)=M(d_k-c_i+a)$.*

It follows from a lemma of Chevalley[5] that a maximal order in $\mathfrak{A}$ has really the form $M(a_{ik})$ (whence the form $M(c_k-c_i)$) whenever it contains all diagonal $\epsilon_{11}, \epsilon_{22}, \cdots, \epsilon_{rr}$.

LEMMA 2. *There exists a regular element α in $\mathfrak{A}$ such that*

$$\alpha^{-1}\mathfrak{J}_1\alpha = M(0), \quad \alpha^{-1}\mathfrak{J}_2\alpha = M(c_k - c_i); \qquad c_1 \geqq c_2 \geqq \cdots \geqq c_r.$$

PROOF. There is, as is well known, a regular element β such that $\beta^{-1}\mathfrak{J}_1\beta=M(0)$. Consider the distance ideal $\mathfrak{d}_{12}=(\mathfrak{J}_2\mathfrak{J}_1)^{-1}=\mathfrak{J}_1\delta$ of $\mathfrak{J}_1$ to $\mathfrak{J}_2$. The theory of elementary divisors tells the existence of two units ξ, η in $M(0)$ such that $\gamma=\xi\beta^{-1}\delta\beta\eta$ is a diagonal matrix with diagonal elements P^{c_i}, $(c_1\geqq\cdots\geqq c_r)$, $\gamma=\sum\epsilon_{ii}P^{c_i}$, where we denote, for the sake of convenience, a prime element of the prime ideal P by the same letter P. Put $\alpha=\beta\eta$. Then this α possesses the required property: $\alpha^{-1}\mathfrak{J}_1\alpha=\eta^{-1}\beta^{-1}\mathfrak{J}_1\beta\eta=M(0)$, $\alpha^{-1}\mathfrak{J}_2\alpha=\gamma^{-1}M(0)\gamma=M(c_k-c_i)$.

LEMMA 3. *There exist two regular elements α, β in $\mathfrak{A}$ such that*

$$\alpha\mathfrak{a}\beta = M(0), \qquad \alpha\mathfrak{b}\beta = M(d_k - c_i);$$

$$c_1 \geqq c_2 \geqq \cdots \geqq c_r, \quad d_1 \geqq d_2 \geqq \cdots \geqq d_r.$$

PROOF. Let $\mathfrak{J}_1'$, $\mathfrak{J}_2'$ $[\mathfrak{J}_3', \mathfrak{J}_4']$ be the left and the right orders of $\mathfrak{a}$ $[\mathfrak{b}]$. According to the above lemma there exist γ, β such that $\gamma^{-1}\mathfrak{J}_1'\gamma=\beta^{-1}\mathfrak{J}_2'\beta=M(0)$, $\gamma^{-1}\mathfrak{J}_3'\gamma=M(c_k-c_i)$, $\beta^{-1}\mathfrak{J}_4'\beta=M(d_k'-d_i')$. $\gamma^{-1}\mathfrak{a}\beta$ is a two-sided ideal of $M(0)$ and has a form $P^aM(0)$. Put $\alpha=(\gamma P^a)^{-1}$. Then $\alpha\mathfrak{a}\beta=M(0)$. Moreover, $\alpha\mathfrak{b}\beta$ is of a form $M(d_k'-c_i+b)$ (Lemma 1). We put $d_k=d_k'+b$, and this completes the proof.

We note further that the left order of an ideal $M(a_{ik})$ is $M(b_{ik})$ where $b_{ik}=\max_j(a_{ij}-a_{kj})$.

After these preliminaries our theorems are easy to prove. In Theorem 1 we may, according to Lemma 2, assume that $\mathfrak{J}_1=M(0)$, $\mathfrak{J}_2=M(c_k-c_i)$, $(c_1\geqq\cdots\geqq c_r)$. Suppose $\mathfrak{J}_1\cap\mathfrak{J}_2=\mathfrak{J}_3\cap\mathfrak{J}_4$. Since

[4] Cf. Nakayama, Japanese Journal of Mathematics, vol. 13 (1937), p. 339.

[5] Chevalley, Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität, vol. 10 (1934), p. 87.

$\epsilon_{ii} \mathfrak{r} \mathfrak{J}_1 \cap \mathfrak{J}_2 \subseteqq \mathfrak{J}_3, \mathfrak{J}_4$, it follows that $\mathfrak{J}_3, \mathfrak{J}_4$ have the form $\mathfrak{J}_3 = M(d_k - d_i)$, $\mathfrak{J}_4 = M(f_k - f_i)$. Moreover max $(d_k - d_i, f_k - f_i) = \max(0, c_k - c_i)$, $(i, k = 1, 2, \cdots, r)$. This implies max $(d_k - d_i, f_k - f_i) = 0$ if $i \geqq k$, whence $d_1 \geqq \cdots \geqq d_r$ and $f_1 \geqq \cdots \geqq f_r$. On applying the same relation to $i = 1$, $k = r$, we find that either $d_1 = d_r$ or $f_1 = f_r$. In the first case we have $d_1 = \cdots = d_r$, $f_1 - f_i = c_1 - c_i$, $(i = 1, 2, \cdots, r)$, whence $\mathfrak{J}_1 = \mathfrak{J}_3$, $\mathfrak{J}_2 = \mathfrak{J}_4$. The second case gives of course $\mathfrak{J}_1 = \mathfrak{J}_4$, $\mathfrak{J}_2 = \mathfrak{J}_3$.

The assertion about the sum follows now from Theorem 2, which is in turn contained in Theorems 3 and 4.

As to Theorem 3 we notice first that if α, β are two regular elements, the ideals $\alpha \mathfrak{r} \alpha^{-1}$, $\beta^{-1} \mathfrak{s} \beta$, $\beta^{-1} \mathfrak{t} \beta$, $\beta^{-1} \mathfrak{t}' \beta$ have the same significance for $\alpha \mathfrak{a} \beta$ and $\alpha \mathfrak{b} \beta$ as the ideals $\mathfrak{r}$, $\mathfrak{s}$, $\mathfrak{t}$, $\mathfrak{t}'$ have for $\mathfrak{a}$ and $\mathfrak{b}$. Hence it is sufficient, by Lemma 3, to consider the case where

$$\mathfrak{a} = M(0), \qquad \mathfrak{b} = M(d_k - c_i); \tag{1}$$
$$c_1 \geqq c_2 \geqq \cdots \geqq c_r, \quad d_1 \geqq d_2 \geqq \cdots \geqq d_r.$$

Then $\mathfrak{a} \cap \mathfrak{b} = M(\max(0, d_k - c_i))$ and $\mathfrak{o} = M(a_{ik})$ with

$$a_{ik} = \max_j (\max(0, d_j - c_i) - \max(0, d_j - c_k))$$
$$= \begin{cases} \max(0, d_1 - c_i) - \max(0, d_1 - c_k) = f_k - f_i & \text{for } i \geqq k, \\ \max(0, d_r - c_i) - \max(0, d_r - c_i) = g_k - g_i & \text{for } i \leqq k, \end{cases}$$

where $f_i = -\max(0, d_1 - c_i)$, $g_i = -\max(0, d_r - c_i)$. Since $f_k - f_i \geqq$ or $\leqq g_k - g_i$ according as $i \geqq k$ or $i \leqq k$, we find that $\mathfrak{o}$ is the intersection of the two maximal orders $M(f_k - f_i)$ and $M(g_k - g_i)$. Further, if we put $\gamma = \sum \epsilon_{ii} P^{-c_i}$, $\delta = \sum \epsilon_{ii} P^{-d_i}$, then $\mathfrak{r} = \gamma P^a M(0)$ and $\mathfrak{s} = M(0) P^{-a} \delta$, whence $\mathfrak{t} = P^{a - d_r} M(0)$, $\mathfrak{t}' = P^{d_1 - a} M(0)$. From this we can easily verify the precise characterization of $\mathfrak{o}$ given in the theorem.

The part on the sum $(\mathfrak{a}, \mathfrak{b})$ can be shown by a similar computation. And indeed from that computation we obtain the last assertion in the theorem.

Finally, to prove Theorem 4 we observe again that we have only to consider the case where $\mathfrak{a}$, $\mathfrak{b}$ have the form (1). $\mathfrak{a} \cap \mathfrak{b} = M(\max(0, d_k - c_i)$, $(\mathfrak{a}^{-1}, \mathfrak{b}^{-1}) = M(\min(0, c_k - d_i))$ because $\mathfrak{b}^{-1} = M(c_k - d_i)$, and here we notice that $\max(0, d_k - c_i) = -\min(0, c_i - d_k)$. The third normal ideal $\mathfrak{c}$ can be expressed as $\mathfrak{c} = \tau^{-1} M(0) \sigma^{-1}$ with regular elements $\sigma = \sum \epsilon_{ik} s_{ik}$, $\tau = \sum \epsilon_{ik} t_{ik}$. Let $P^{c_{ik}}$ be the exact power of P which divides s_{ik}, $P^{c_{ik}} \| s_{ik}$; if $s_{ik} = 0$ we put $c_{ik} = \infty$. Let similarly $P^{d_{ik}} \| t_{ik}$. It is evident that $M(a_{ik})$, with a system of rational integers a_{ik}, contains $\mathfrak{c}^{-1} = \sigma M(0) \tau$ if and only if

$$c_{ij} + d_{lk} \geqq a_{ik}, \qquad \text{for all } i, j, k, l. \tag{2}$$

Hence, if we show that the same condition is also necessary and sufficient in order that $M(-a_{ki}) \subseteqq \mathfrak{c}$, then we will be through. But this is also easy to see. For, $\mathfrak{c}=\tau^{-1}M(0)\sigma^{-1}$ consists of all $\eta=\sum \epsilon_{ik}y_{ik}$ $=\tau^{-1}(\sum \epsilon_{ik}x_{ik})\sigma^{-1}$ with $x_{ik} \in I$. On taking a pair (j, l) of indices, let us consider those η such that $y_{ik}=0$ for $(i, k)\neq(j, l)$. In other words, we consider the equation $\tau^{-1}(\sum \epsilon_{ik}x_{ik})\sigma^{-1}=\epsilon_{jl}y_{jl}$. But this is equivalent to $\sum \epsilon_{ik}x_{ik}=\tau\epsilon_{jl}y_{jl}\sigma$, or

$$x_{ik} = t_{ij}y_{jl}s_{lk}, \qquad i, k = 1, 2, \cdots, r. \tag{3}$$

Suppose now $M(-a_{ki}) \subseteqq \mathfrak{c}$. Then (3) with $y_{jl}=P^{-a_{lj}}$ must have a solution $x_{ik} \in I$. Hence $0 \leqq d_{ij}-a_{lj}+c_{lk}$ (for all i, k). Since (j, l) was an arbitrary pair of indices, we have thus established (2). Assume conversely (2). Then obviously $x_{ik}=t_{ij}P^{-a_{lj}}s_{lk} \in I$ whence $\epsilon_{jl}P^{-a_{lj}} \in \mathfrak{c}$ and $M(-a_{ki}) \subseteqq \mathfrak{c}$.

A second proof of the last part of Theorem 3 is as follows: We observe first that every ideal $\mathfrak{m}$ in $\mathfrak{A}$ is additively generated by regular elements contained in $\mathfrak{m}$.[6] For, if $\xi \in \mathfrak{m}$ we take a scalar element a $(\in F)$ in $\mathfrak{m}$ different from all the characteristic roots of the matrix which represents ξ in a faithful representation of $\mathfrak{A}$. Then $\xi-a$ $(\in \mathfrak{m})$ is evidently a regular element and $\xi=(\xi-a)+a$. Now, let α be any regular element from the left order of $\mathfrak{a} \cap \mathfrak{b}$; $\alpha(\mathfrak{a} \cap \mathfrak{b}) \subseteqq \mathfrak{a} \cap \mathfrak{b}$. Since $\alpha\mathfrak{a}$ and $\alpha\mathfrak{b}$ are normal ideals, we have, from Theorem 4, $(\mathfrak{a}^{-1}\alpha^{-1}, \mathfrak{b}^{-1}\alpha^{-1}) \supseteqq \mathfrak{a}^{-1}$, $\mathfrak{b}^{-1}$ whence $(\mathfrak{a}^{-1}, \mathfrak{b}^{-1})\alpha^{-1} \supseteqq (\mathfrak{a}^{-1}, \mathfrak{b}^{-1})$, $(\mathfrak{a}^{-1}, \mathfrak{b}^{-1}) \supseteqq (\mathfrak{a}^{-1}, \mathfrak{b}^{-1})\alpha$. This shows that the left order of $\mathfrak{a} \cap \mathfrak{b}$ is contained in the right order of $(\mathfrak{a}^{-1}, \mathfrak{b}^{-1})$. But the converse can be seen in quite a similar manner.

Remark. The structure of the residue class algebra $\mathfrak{J}_1 \cap \mathfrak{J}_2/p(\mathfrak{J}_1 \cap \mathfrak{J}_2)$ is easy to analyze, but perhaps does not deserve a detailed discussion. We merely note that the algebra is not symmetric, in fact is not weakly symmetric,[7] except for the trivial case $(\mathfrak{J}_1)_p=(\mathfrak{J}_2)_p$; this remark may be of some interest in view of a recent paper by R. Brauer.[8]

THE INSTITUTE FOR ADVANCED STUDY

[6] We exclude here the trivial case of a finite underlying field F.

[7] See Brauer-Nesbitt, Proceedings of the National Academy of Sciences, vol. 23 (1937); Nakayama-Nesbitt, Annals of Mathematics, (2), vol. 39 (1938).

[8] Brauer, *On modular and p-adic representations of algebras*, Proceedings of the National Academy of Sciences, vol. 25 (1939).

A NOTE ON CONVEX FUNCTIONS[1]

BERNARD FRIEDMAN

It is well known that, just from the definition of convexity, a function of one variable which is convex must also be absolutely continuous. It was not known whether this property would generalize to convex functions of more than one variable. In this note there is constructed a convex function of two variables which is not absolutely continuous.

Carathéodory's definition of absolute continuity [1] for functions of two variables may be stated as follows:

Let $f(x, y)$ be any functions of two variables and S the square with the vertices

$$(x - h, y - h), \quad (x - h, y + h), \quad (x + h, y - h), \quad (x + h, y + h).$$

Let $I(S)$ be the following interval function

$$I(S) = f(x + h, y + h) + f(x - h, y - h) - f(x + h, y - h) - f(x - h, y + h).$$

Then $f(x, y)$ is said to be absolutely continuous in a two-dimensional region R if $f(x, y)$ is absolutely continuous in each variable separately and if also $I(S)$ is an absolutely continuous function of S. Carathéodory [2] proves the absolute continuity of $f(x, y)$ is a necessary and sufficient condition that there exist functions $g(x)$, $h(y)$ and $k(x, y)$ such that

$$f(x, y) = \int_0^x g(x)dx + \int_0^y h(y)dy + \int_0^x \int_0^y k(x, y)dxdy + C.$$

The counterexample of this note will be constructed from the integral of Cantor's middle third function which can be defined as follows:

$$\begin{aligned} m(0) &= 0, & m(1) &= 1, \\ m(x) &= 1/2, & 1/3 &\leqq x \leqq 2/3, \\ m(x) &= 1/4, & 1/9 &\leqq x \leqq 2/9, \\ m(x) &= 3/4, & 7/8 &\leqq x \leqq 8/9, \cdots . \end{aligned}$$

It is well known that $m(x)$ is a continuous monotonically increasing function whose derivative is zero almost everywhere. Obviously $m(x)$

[1] Presented to the Society, April 8, 1938.

is not absolutely continuous and not convex, but $f(x)=\int_0^x m(t)dt$ will be both convex and absolutely continuous. $f(x)$ is convex since its derivative $m(x)$ is everywhere non-decreasing.

It is to be noticed that if $f(x)$ is a convex function of x, then $f(x, y)\equiv f(x+y)$ is a convex function of two variables x and y.

However if $f(x, y)$ were absolutely continuous in the two variables x and y, by Carathéodory's theorem it could be represented as

$$f(x+y)=\int_0^x g(x)dx+\int_0^y h(y)dy+\int_0^x\int_0^y k(x, y)dxdy+C.$$

But differentiating with respect to x it follows that

$$f'(x+y)=m(x+y)=g(x)+\int_0^y k(x, y)dy$$

for almost all values of x. Let x_0 be some value of x for which the equality holds. Then

$$m(x_0+y)=g(x_0)+\int_0^y k(x_0, y)dy$$

or $m(x_0+y)$ would be absolutely continuous, which is impossible.

Therefore $f(x+y)$ is a convex function of two variables which is not absolutely continuous.

It is quite easy to prove directly that $I(S)$ is not absolutely continuous. The direct proof, however, is uninteresting and since it requires some tedious computations, it will be omitted.

Bibliography

1. Carathéodory, *Vorlesungen über Reele Funktionen*, 1918, p. 653.
2. Carathéodory, loc. cit., p. 654.

University of Wisconsin

REAL ASYMPTOTIC SOLUTIONS OF REAL DIFFERENTIAL EQUATIONS[1]

MONROE H. MARTIN

Let us consider the system of differential equations

$$(1) \qquad \dot{x}_k = \sum_{j=1}^{n} a_{kj}x_j + \phi_k(x_1, \cdots, x_n), \qquad \dot{x}_k = \frac{dx_k}{dt}, \qquad k = 1, \cdots, n,$$

in which the a_{kj} denote constants and the $\phi_k(x_1, \cdots, x_n)$ are power series in $x_1, \cdots, x_n$ converging in a given neighborhood of the origin $x_1 = \cdots = x_n = 0$, and containing neither constant nor linear terms. The differential equations in the system are supposed to be real and we shall investigate real solutions

$$(2) \qquad x_k = x_k(t), \qquad k = 1, \cdots, n,$$

which are asymptotic to the origin as t tends to $+\infty$ through real values, that is, real solutions for which

$$(3) \qquad \lim_{t \to +\infty} (x_1^2 + \cdots + x_n^2) = 0.$$

When $n = 2$ the results to which we shall be led are well known[2] and may be formulated as follows:

If the characteristic constants of the matrix $\|a_{kj}\|$ *in* (1) *are both negative and one is not an integral multiple of the other, or are conjugate complex numbers with negative real parts, all the points in a suitably restricted neighborhood of the origin lie on real solutions of* (1) *which are asymptotic to the origin as* t *tends to* $+\infty$. *If the characteristic constants are real and of opposite signs, a point in a suitably restricted neighborhood of the origin lies on a real solution of* (1) *asymptotic to the origin as* t *tends to* $+\infty$ *if, and only if, it is a point of a certain real analytic arc containing the origin.*

Recent dynamical investigations[3] by the author require an exten-

[1] Presented to the Society, December 29, 1938, under the title *Restricted problems in three bodies.*

[2] See, for example, E. Picard, *Traité d'Analyse*, vol. 3, 1928, pp. 206–213. In connection with the case where the characteristic constants have opposite signs, cf. P. Painlevé, *Gewöhnliche Differentialgleichungen; Existenz der Lösungen*, Encyklopädie der mathematischen Wissenschaften, vol. 2, pt. 1^1, p. 221 (footnote 116), where it is pointed out that Poincaré failed to recognize the necessity of proving what is equivalent to showing that a point not on the stated analytic arc cannot lie on a solution of (1) asymptotic to the origin as t tends to $+\infty$. See also E. Picard, loc. cit., p. 207.

[3] Not yet published.

sion of this theorem to the cases $n>2$. That the extension is not obvious has been maintained by Bieberbach.[4] The difficulty in generalizing the theorem arises when some of the characteristic constants of the matrix $\|a_{kj}\|$ are complex. The need for a proof of the existence of real solutions asymptotic to the origin when some of the characteristic constants are complex has been emphasized by Wintner.[5]

The general theorem to be established follows.

THEOREM. *If the first m of the characteristic constants λ_k of the matrix $\|a_{kj}\|$ in the system* (1) *are negative or have negative real parts, while the remaining $n-m$ are positive or have positive real parts, if the elementary divisors of the characteristic matrix are linear, and if none of the linear commensurability relations*

$$\lambda_k = p_1\lambda_1 + \cdots + p_m\lambda_m,$$

$$k = 1, \cdots, m;\ p_k = 0, 1, \cdots;\ p_1 + \cdots + p_m \geqq 2,$$

holds, a point in a suitably restricted neighborhood of the origin lies on a real solution of (1) *asymptotic to the origin as t tends to $+\infty$ if, and only if, it is a point of a certain real, analytic, m-dimensional manifold regular at the origin.*

A companion theorem corresponding to the remaining $n-m$ characteristic constants and dealing with real solutions asymptotic to the origin as t tends to $-\infty$ may obviously be formulated.

Our investigation is based on the following results[6] due to Poincaré, Picard and Horn.

Assume given the system of differential equations

$$\dot{z}_k = \lambda_k z_k + \phi_k(z_1, \cdots, z_n), \qquad k = 1, \cdots, n,$$

in which z_k denote real or complex variables, $\phi_k(z_1, \cdots, z_n)$ denote real or complex power series in $z_1, \cdots, z_n$, which converge in a given neighborhood of $z_1 = \cdots = z_n = 0$ and begin with terms of at least the second degree in $z_1, \cdots, z_n$. The constants λ_k may be real or complex and it is assumed that $\lambda_1, \cdots, \lambda_m$ have negative real parts while $\lambda_{m+1}, \cdots, \lambda_n$ have positive real parts. In addition the constants λ_k are subject to the hypothesis that none of the linear commensurability relations

$$\lambda_k = p_1\lambda_1 + \cdots + p_m\lambda_m, \qquad k = 1, \cdots, m,$$

[4] L. Bieberbach, *Differentialgleichungen*, 3d edition, 1930, p. 113.

[5] A. Wintner, *Librationstheorie des restringierten Dreikörperproblems*, Mathematische Zeitschrift, vol. 32 (1930), pp. 661–662.

[6] E. Picard, op. cit., pp. 17–21, and pp. 187–188; J. Horn, *Gewöhnliche Differentialgleichungen beliebiger Ordnung*, 1905, pp. 319–325. For (5) in the text above, see page 321 of Horn's book.

holds, where $p_1, \cdots, p_m$ *denote non-negative integers for which* $p_1 + \cdots + p_m \geqq 2$.

If $C_1, \cdots, C_m$ *denote m arbitrary real or complex constants, it is possible to find power series for* $z_1, \cdots, z_n$ *in the arguments*

$$u_1 = C_1 e^{\lambda_1 t}, \cdots, u_m = C_m e^{\lambda_m t}, \tag{4}$$

which converge for sufficiently small absolute values of these arguments and which satisfy the differential equations. These power series have the form

$$\begin{aligned} z_j &= u_j + P_j(u_1, \cdots, u_m), \qquad j = 1, \cdots, m, \\ z_k &= P_k(u_1, \cdots, u_m), \qquad k = m+1, \cdots, n, \end{aligned} \tag{5}$$

where the power series $P_1, \cdots, P_n$ *begin with terms of at least the second degree in* $u_1, \cdots, u_m$.

The complex characteristic constants of the real matrix $\|a_{kj}\|$ necessarily occur in conjugate pairs and we arrange the notation so that

$$\begin{aligned} \bar{\lambda}_{2\nu_1-1} &= \lambda_{2\nu_1}, & \nu_1 &= 1, \cdots, j_0, \\ \bar{\lambda}_{\nu_2} &= \lambda_{\nu_2} < 0, & \nu_2 &= 2j_0+1, \cdots, m, \\ \bar{\lambda}_{m+2\nu_3-1} &= \lambda_{m+2\nu_3}, & \nu_3 &= 1, \cdots, k_0, \\ \bar{\lambda}_{\nu_4} &= \lambda_{\nu_4} > 0, & \nu_4 &= m+2k_0+1, \cdots, n, \end{aligned} \tag{6}$$

where j_0, k_0 denote integers such that $0 \leqq 2j_0 \leqq m$, $0 \leqq 2k_0 \leqq n-m$. The ν_1, ν_2, ν_3, ν_4 will always be used in the sequel in the same sense as employed here. Thus, to indicate the real and imaginary parts of the complex characteristic constants, we shall write

$$\begin{aligned} \lambda_{2\nu_1-1} &= \alpha_{\nu_1} + i\beta_{\nu_1}, & \lambda_{2\nu_1} &= \alpha_{\nu_1} - i\beta_{\nu_1}, \\ \lambda_{m+2\nu_3-1} &= \alpha_{m+\nu_3} + i\beta_{m+\nu_3}, & \lambda_{m+2\nu_3} &= \alpha_{m+\nu_3} - i\beta_{m+\nu_3}. \end{aligned}$$

From the theory of linear transformations it follows that there exists a real linear transformation with nonvanishing determinant transforming the variables $x_1, \cdots, x_n$ into the variables $y_1, \cdots, y_n$ and the system (1) of differential equations into the system

$$\begin{aligned} \dot{y}_{2\nu_1-1} &= \alpha_{\nu_1} y_{2\nu_1-1} - \beta_{\nu_1} y_{2\nu_1} + \Phi_{2\nu_1-1}(y_1, \cdots, y_n), \\ \dot{y}_{2\nu_1} &= \beta_{\nu_1} y_{2\nu_1-1} + \alpha_{\nu_1} y_{2\nu_1} + \Phi_{2\nu_1}(y_1, \cdots, y_n), \\ \dot{y}_{\nu_2} &= \lambda_{\nu_2} y_{\nu_2} + \Phi_{\nu_2}(y_1, \cdots, y_n), \\ \dot{y}_{m+2\nu_3-1} &= \alpha_{m+\nu_3} y_{m+2\nu_3-1} - \beta_{m+\nu_3} y_{m+2\nu_3} + \Phi_{m+2\nu_3-1}(y_1, \cdots, y_n), \\ \dot{y}_{m+2\nu_3} &= \beta_{m+\nu_3} y_{m+2\nu_3-1} + \alpha_{m+\nu_3} y_{m+2\nu_3} + \Phi_{m+2\nu_3}(y_1, \cdots, y_n), \\ \dot{y}_{\nu_4} &= \lambda_{\nu_4} y_{\nu_4} + \Phi_{\nu_4}(y_1, \cdots, y_n), \end{aligned} \tag{7}$$

where $\Phi_1, \cdots, \Phi_n$ denote real power series in $y_1, \cdots, y_n$ beginning with terms of at least the second degree in these variables.

If complex variables $z_1, \cdots, z_n$ be introduced by means of the equations

$$(8)\quad \begin{aligned} z_{2\nu_1-1} &= y_{2\nu_1-1} + iy_{2\nu_1}, \qquad z_{2\nu_1} = y_{2\nu_1-1} - iy_{2\nu_1}, \qquad z_{\nu_2} = y_{\nu_2}, \\ z_{m+2\nu_3-1} &= y_{m+2\nu_3-1} + iy_{m+2\nu_3}, z_{m+2\nu_3} = y_{m+2\nu_3-1} - iy_{m+2\nu_3}, \ z_{\nu_4} = y_{\nu_4}, \end{aligned}$$

the differential equations (7) take the simple form

$$(9)\quad \dot{z}_k = \lambda_k z_k + \sum_{p_1+\cdots+p_n\geqq 2} A^{(k)}_{p_1\cdots p_n} z_1^{p_1} \cdots z_n^{p_n}, \qquad k = 1, \cdots, n,$$

where

$$(10)\quad \overline{A}^{(k)}_{p_1\cdots p_n} = A^{(j)}_{q_1\cdots q_n} \qquad \text{whenever } \overline{\lambda}_k = \lambda_j,$$

and $q_1, \cdots, q_n$ is the permutation of $p_1, \cdots, p_n$ obtained by interchanging every pair p_k, p_l for which k and l are such that $\overline{\lambda}_k = \lambda_l$.

It is apparent from (8) that a solution of (7) will be real if, and only if, the corresponding solution of (9) satisfies the condition

$$(11)\quad \bar{z}_k = z_j \qquad \text{whenever } \overline{\lambda}_k = \lambda_j,$$

and we shall now construct solutions of (9) satisfying these conditions.

The following method was suggested by my colleague E. Titt and I have adopted it in preference to my own as less cumbersome.

Referring to the definition of $\nu_1, \nu_2, \nu_3, \nu_4$ in (6), it follows from (10) that if

$$(12)\quad z_{2\nu_1-1}, z_{2\nu_1}, z_{\nu_2}, z_{m+2\nu_3-1}, z_{m+2\nu_3}, z_{\nu_4}$$

denotes a set of n real or complex functions of t which is a solution of (9), then the set

$$(13)\quad \bar{z}_{2\nu_1}, \bar{z}_{2\nu_1-1}, \bar{z}_{\nu_2}, \bar{z}_{m+2\nu_3}, \bar{z}_{m+2\nu_3-1}, \bar{z}_{\nu_4},$$

obtained from (12) by replacing $z_k(t)$ by $\bar{z}_j(t)$ whenever $\overline{\lambda}_k = \lambda_j$ $(1 \leqq k \leqq j \leqq n)$, also represents a solution of (9). Accordingly a solution of (7) obtained from (8), will be real if, and only if, the two solutions (12) and (13) of (9) are identical.

The results of Poincaré and Picard apply to the system (9) and we write the solution (5) in the form

$$(14)\quad z_k = \sum_{k_1+\cdots+k_m\geqq 1} C^{(k)}_{k_1\cdots k_m} u_1^{k_1} \cdots u_m^{k_m},$$

where we have, of course, when $k_1+\cdots+k_m=1$,

$$(15)\quad C^{(j)}_{k_1\cdots k_m}=1,\quad k_1=0,\cdots,k_j=1,\cdots,k_m=0;j=1,\cdots,m,$$

the $C^{(k)}_{k_1\cdots k_m}$ remaining, for which $k_1+\cdots+k_n=1$, $k=1,\cdots,n$, being zero. Choose the arbitrary constants $C_1,\cdots,C_m$ in (4) so that

$$(16)\qquad \overline{C}_p=C_q \qquad \text{whenever } \overline{\lambda}_p=\lambda_q,\ (p\leqq q\leqq m),$$

and therefore

$$(17)\qquad \bar{u}_p=u_q \qquad \text{whenever } \overline{\lambda}_p=\lambda_q,\ (p\leqq q\leqq m).$$

Selecting the solution (14) for the solution in (12), we shall prove that it is identical with the solution (13). Upon substituting (14) in (9) and equating coefficients of $u_1^{k_1}\cdots u_m^{k_m}$ it may be verified that $C^{(k)}_{k_1\cdots k_m}$, for $k_1+\cdots+k_m=r$ $(r\geqq 2)$ and k fixed, is expressible as a polynomial in the $C^{(k)}_{k_1\cdots k_m}$ with $k_1+\cdots+k_m<r$ and $k=1,2,\cdots,n$. These recurrence equations determine the $C^{(k)}_{k_1\cdots k_m}$ uniquely, once the coefficients of the linear terms in $u_1,\cdots,u_n$ of (14) are given. Taking the coefficients of the linear terms in $u_1,\cdots,u_m$ of (12) as given by (15), it follows from (17) that they are identical with the coefficients of the linear terms in $u_1,\cdots,u_m$ of (13), and the two solutions are therefore identical.

Accordingly, the constants C_j in (4) being chosen subject to (16) and the complex solution (14) of (9) inserted in (8), a real solution of (7) is obtained when equations (8) are solved for $y_1,\cdots,y_n$.

If we write

$$(18)\qquad u_{2\nu_1-1}=v_{2\nu_1-1}+iv_{2\nu_1},\qquad u_{2\nu_1}=v_{2\nu_1-1}-iv_{2\nu_1},\qquad u_{\nu_2}=v_{\nu_2},$$

the afore-mentioned real solution of (7) takes the form

$$(19)\qquad \begin{aligned} y_j&=v_j+Q_j(v_1,\cdots,v_m), & j&=1,\cdots,m,\\ y_k&=Q_k(v_1,\cdots,v_m), & k&=m+1,\cdots,n,\end{aligned}$$

where $Q_1,\cdots,Q_n$ denote real power series beginning with terms of at least the second degree in the real variables $v_1,\cdots,v_m$, and converging for sufficiently small $|v_1|,\cdots,|v_m|$. If we let

$$C_{2\nu_1-1}=a_{\nu_1}+ib_{\nu_1},\qquad C_{2\nu_1}=a_{\nu_1}-ib_{\nu_1},\qquad C_{\nu_2}=a_{\nu_2},$$

it is readily seen that

$$(20)\qquad \begin{aligned} v_{2\nu_1-1}&=e^{\alpha\nu_1 t}(a_{\nu_1}\cos\beta_{\nu_1}t-b_{\nu_1}\sin\beta_{\nu_1}t),\\ v_{2\nu_1}&=e^{\alpha\nu_1 t}(a_{\nu_1}\sin\beta_{\nu_1}t+b_{\nu_1}\cos\beta_{\nu_1}t),\qquad v_{\nu_2}=a_{\nu_2}e^{\lambda\nu_2 t}.\end{aligned}$$

The constants a_{r_1}, b_{r_1}, a_{r_2} are real and arbitrary. It therefore follows from (19) and (20) that there exists a real, analytic, m-dimensional manifold of solutions of (7) in a sufficiently small neighborhood of the origin $y_1 = \cdots = y_n = 0$. This m-dimensional manifold is regular at the origin, and, since $\alpha_{r_1}, \lambda_{r_2}$ are negative, any one of its points lies on a real solution of (1) asymptotic to the origin as t tends to $+\infty$.

We shall now show that a sufficiently small neighborhood of the origin $y_1 = \cdots = y_n = 0$ contains no points on real solutions of (7) asymptotic to the origin as t tends to $+\infty$ other than the points of the above manifold.

Solving the first m equations in (19) for $v_1, \cdots, v_m$ and substituting these expressions in the remaining $n-m$ equations, we find the equations of the m-dimensional manifold in non-parametric form, namely

$$y_k = Q_k^*(y_1, \cdots, y_m), \qquad k = m+1, \cdots, n, \tag{21}$$

where Q_k^* denote power series beginning with terms of the second degree at least in $y_1, \cdots, y_m$ and converging for sufficiently small $|y_1|, \cdots, |y_m|$.

Let us introduce new coordinates $\eta_1, \cdots, \eta_n$ by means of the equations

$$\eta_j = y_j, \qquad j = 1, \cdots, m,$$

$$\eta_k = y_k - Q_k^*(y_1, \cdots, y_m), \qquad k = m+1, \cdots, n.$$

The differential equations (7) retain their form after transformation, the y's being replaced, of course, by η's. The coefficients of the linear terms are unaltered and the power series $\Phi_1, \cdots, \Phi_n$ become power series in $\eta_1, \cdots, \eta_n$ which will be denoted by $\Phi_1^*, \cdots, \Phi_n^*$ respectively.

The last $n-m$ of the transformed equations must be satisfied when we put $\eta_{m+1} = \cdots = \eta_n = 0$, and therefore

$$\Phi_k^* = \sum_{l=m+1}^{n} \Phi_{kl}\eta_l, \qquad k = m+1, \cdots, n, \tag{22}$$

where Φ_{kl} denote power series in $\eta_1, \cdots, \eta_n$ without constant terms. Multiplying the last $n-m$ transformed equations by $\eta_{m+1}, \cdots, \eta_n$, respectively and adding, there results

$$\sum_{k=m+1}^{n} \eta_k\dot{\eta}_k = \sum_{k=m+1}^{n} \alpha_k\eta_k^2 + \sum_{k=m+1}^{n}\sum_{l=m+1}^{n} \Phi_{kl}\eta_k\eta_l,$$

where we have employed (22) to substitute for Φ_k^*. Writing

$$\rho^2 = \eta_{m+1}^2 + \cdots + \eta_n^2, \qquad \mu_k = \eta_k/\rho, \qquad k = m+1, \cdots, n,$$

in the above equation and dividing by ρ^2, we find

$$\dot{\rho}/\rho = \sum_{k=m+1}^{n} \alpha_k \mu_k^2 + \sum_{k=m+1}^{n} \sum_{l=m+1}^{n} \Phi_{kl} \mu_k \mu_l. \tag{23}$$

Let $\alpha^* = \min \{\alpha_{m+1}, \cdots, \alpha_n\}$ and $\Phi^* = \max \{|\Phi_{kl}|\}$, where $\Phi^* = \Phi^*(\eta_1, \cdots, \eta_n)$ tends to zero as the point $(\eta_1, \cdots, \eta_n)$ tends towards the origin $\eta_1 = \cdots = \eta_n = 0$. Now $\mu_{m+1}^2 + \cdots + \mu_n^2 = 1$, so that we have

$$\left| \sum_{k=m+1}^{n} \sum_{l=m+1}^{n} \Phi_{kl} \mu_k \mu_l \right| \leqq (n-m)\Phi^*,$$

and hence from (23) we see that

$$\dot{\rho}/\rho \geqq \alpha^* - (n-m)\Phi^*, \qquad \alpha^* > 0. \tag{24}$$

If possible let $y_k = y_k(t)$ be a real solution of (7) which is asymptotic to the origin $y_1 = \cdots = y_n = 0$ as t tends to $+\infty$, yet for which there is no $\bar{t}$ such that the solution lies in the m-dimensional manifold (21) for $t > \bar{t}$.

Employing $\eta_1, \cdots, \eta_n$ as coordinates, we should then have a real solution $\eta_k = \eta_k(t)$ of the transformed differential equations with the property that to any positive number ϵ there corresponds a $t_\epsilon > 0$ such that we have

$$\eta_1^2 + \cdots + \eta_n^2 < \epsilon \tag{25}$$

for all $t > t_\epsilon$. In addition there must exist a $t_0 > t_\epsilon$ which is such that on writing $\eta_k^0 = \eta_k(t_0)$, $\rho_0^2 = \eta_{m+1}^{0^2} + \cdots + \eta_n^{0^2}$ $(k = 1, \cdots, n)$, we have $\rho_0 > 0$. Since Φ^* is arbitrarily small in the neighborhood of the origin of $\eta_1 = \cdots = \eta_n = 0$, it is possible to choose ϵ at the outset so small that

$$\Phi^* = \Phi^*(\eta_1(t), \cdots, \eta_n(t)) < \alpha^*/2(n-m) \qquad \text{for } t > t_\epsilon.$$

Inequality (24) then becomes $\dot{\rho}/\rho > \alpha^*/2$ for $t > t_\epsilon$. Integrating from t_0 to t $(t > t_0)$, we find

$$\rho > \rho_0 e^{(\alpha^*/2)(t-t_0)},$$

which, since $\alpha^* > 0$, implies a contradiction to (25).

Since the transformation from the system (1) to the system (7) was a real linear transformation with nonvanishing determinant the theorem has now been established.

UNIVERSITY OF MARYLAND

ON THE SUPPORTING-PLANE PROPERTY OF A CONVEX BODY[1]

DAVID MOSKOVITZ AND L. L. DINES

In an earlier paper,[2] the authors have shown that in a linear space $\mathfrak{S}$ with an inner product, a set $\mathfrak{M}$ which is closed and linearly connected is supported at a set of boundary points which is everywhere dense on the boundary of $\mathfrak{M}$, and an example is given to show that such a set $\mathfrak{M}$ may have boundary points through which no supporting plane exists. The purpose of this paper is to show that if a set, in addition to being linearly connected and closed, also possesses inner points, then it is completely supported at its boundary points. In (I), reference was made to a paper by Ascoli in which such a result was obtained in a separable space. We do not assume our space $\mathfrak{S}$ to be separable. The definitions and results of (I) will be used in this paper.

A set $\mathfrak{K}$, which is a proper subset of the space $\mathfrak{S}$, will be called a *convex body* if it is linearly connected, closed, and possesses inner points. In the sequel $\mathfrak{K}$ will always denote a convex body.

With reference to the set $\mathfrak{K}$, there is associated with each point x of the space $\mathfrak{S}$ a nonnegative number $r(x)$: if x is an inner point of $\mathfrak{K}$, $r(x)$ is defined as the least upper bound of the radii of spheres about x which do not contain points exterior to $\mathfrak{K}$; for other points of $\mathfrak{S}$, $r(x)$ is defined to be zero. We will call $r(x)$ *the radius at the point* x.

If x_1 is a point of $\mathfrak{K}$, all points x of the sphere $\|x-x_1\| \leqq r(x_1)$ are points of $\mathfrak{K}$.

THEOREM 1. *Let r_1 and r_2 be the radii at the points x_1 and x_2, respectively, of the convex body $\mathfrak{K}$. Then the radius r at the point*

$$x = x_1 + k(x_2 - x_1), \qquad 0 \leqq k \leqq 1,$$

satisfies

$$r \geqq r_1 + k(r_2 - r_1).$$

PROOF. Let $y=x+\rho u$, where $\rho=r_1+k(r_2-r_1)$ and $\|u\|=1$. The points $y_1=x_1+r_1u$ and $y_2=x_2+r_2u$ are points of $\mathfrak{K}$. But from the definitions of x, ρ, and y, it follows that $y=y_1+k(y_2-y_1)$. Hence y, being on the segment joining y_1 and y_2, is also a point of $\mathfrak{K}$. Consequently

[1] Presented to the Society, September 5, 1939.

[2] *On convexity in a linear space with an inner product*, Duke Mathematical Journal, vol. 5 (1939), pp. 520–534. Hereafter, this paper will be referred to by (I).

all points on the boundary of the sphere with radius ρ and center x are in $\mathfrak{K}$. Since $\mathfrak{K}$ is linearly connected, also all points within this sphere are in $\mathfrak{K}$. Therefore $r \geqq \rho$ and the theorem is established.

The following corollaries, which appear self-evident in ordinary space, can be shown to be direct consequences of the preceding theorem.

COROLLARY 1. *Each point of the segment joining two inner points of $\mathfrak{K}$ is an inner point of $\mathfrak{K}$.*

COROLLARY 2. *If x_0 is a boundary point and x_1 an inner point of $\mathfrak{K}$, then the points $x = x_1 + k(x_0 - x_1)$ are inner points of $\mathfrak{K}$ for $0 \leqq k < 1$, and exterior points for $k > 1$.*

With reference to a given boundary point x_0 of the set $\mathfrak{K}$, there is associated with each point x, other than x_0, of the space $\mathfrak{S}$ a non-negative number $s(x)$, defined by[3]

$$s(x) = r(x)/\|x - x_0\|.$$

If x is an exterior point or a boundary point of $\mathfrak{K}$, other than x_0, $s(x)$ is equal to zero; if x is an inner point of $\mathfrak{K}$, $s(x)$ is positive; $s(x_0)$ is not defined.

It is also obvious that $s(x) \leqq 1$, since $r(x) \leqq \|x - x_0\|$.

THEOREM 2. *Let x_0 be a given boundary point of the convex body $\mathfrak{K}$, and let x_t be given by*

$$x_t = x_0 + tu, \qquad \text{where } t > 0, \ \|u\| = 1. \tag{1}$$

Then, for fixed u,

(a) *$s(x_t)$ is a non-decreasing function as $t \to 0$; and*

(b) *$\lim_{t \to 0} s(x_t)$ exists.*

PROOF. In case there are no points of $\mathfrak{K}$ given by (1), the theorem is obviously true, for then

$$s(x_t) = 0 \quad \text{for} \quad t > 0, \qquad \lim_{t \to 0} s(x_t) = 0.$$

In case there are points of $\mathfrak{K}$ given by (1), let x_1 and x_2 be two points of $\mathfrak{K}$ on (1) for parameter values t_1 and t_2, where $t_1 < t_2$; then we have

$$\begin{aligned} x_1 &= x_0 + t_1 u, & x_2 &= x_0 + t_2 u; \\ s(x_1) &= r(x_1)/t_1, & s(x_2) &= r(x_2)/t_2. \end{aligned} \tag{2}$$

[3] Since s is a function of x_0 as well as x, a more explicit notation would be $s(x_0, x)$; but the simpler notation will suffice, inasmuch as the function is to be used in the sequel only with reference to a fixed boundary point x_0.

But $x_1=x_0+(t_1/t_2)(x_2-x_0)$, and hence, by Theorem 1, we have

$$r(x_1) \geqq \frac{t_1}{t_2} r(x_2),$$

since $r(x_0)=0$. Therefore, by (2), $s(x_1)\geqq s(x_2)$.

This result establishes part (a) of the theorem. Since $s(x_t)$ cannot exceed one, obviously part (b) of the theorem is true.

Let Σ be the unit sphere about x_0; and let p_u be the point on Σ given by $p_u=x_0+u$, $\|u\|=1$. Let $x_t=x_0+tu$, $(0<t<1)$, be the segment joining x_0 to p_u; and let[4]

$$\sigma(u) = \lim_{t\to 0} s(x_t).$$

We thus have a function $\sigma(u)$ uniquely defined at each point p_u on the sphere Σ. Obviously, by its definition, we have

$$0 \leqq \sigma(u) \leqq 1.$$

Also $\sigma(u)=0$ only if the segment joining x_0 to p_u does not contain any inner points of $\mathfrak{R}$. If the segment joining x_0 to p_u contains inner points of $\mathfrak{R}$, we have $\sigma(u)>0$.

LEMMA 1. *Let p_u and p_v be two points on Σ, such that*

$$p_u = x_0 + u, \qquad p_v = x_0 + v, \qquad v = -u.$$

Then at least one of the numbers $\sigma(u)$ or $\sigma(v)$ is equal to zero.

PROOF. Assume $\sigma(u)>0$; then the segment joining x_0 to p_u contains inner points. Consequently, by Corollary 2, the segment joining x_0 to p_v does not contain any inner points. Therefore, $\sigma(v)=0$.

THEOREM 3. *Let x_0 be a given boundary point of the convex body $\mathfrak{R}$, and let Σ be the unit sphere about x_0. Let p_u and p_v given by*

$$p_u = x_0 + u, \qquad \|u\| = 1, \qquad p_v = x_0 + v, \qquad \|v\| = 1$$

be two distinct points on Σ, for which $\sigma(u)$ and $\sigma(v)$ are both positive. Then there exists a point p_w distinct from p_u and p_v for which

$$\sigma(w) > \tfrac{1}{2}[\sigma(u) + \sigma(v)].$$

---

[4] The limit was shown to exist in Theorem 2; we are denoting the value of this limit by $\sigma(u)$. It may be of interest to note that

$$\sigma(u) = \lim_{x\to x_0, \text{ along } x_t = x_0+tu} s(x) = \lim_{x\to x_0} \frac{r(x)-r(x_0)}{\|x-x_0\|} = r'_u(x_0)$$

is the *directional derivative* of $r(x)$ at x_0 in the direction u.

PROOF. Let x_t and y_t be points of $\mathfrak{K}$ given by

$$x_t = x_0 + tu, \qquad y_t = x_0 + tv, \qquad 0 < t < 1,$$

and let $((u, v)) = \lambda$. Then, certainly $|\lambda| \leqq 1$. But if $\lambda = 1$, $u = v$, and p_u and p_v are not distinct. If $\lambda = -1$, $u = -v$, in which case not both of the numbers $\sigma(u)$ and $\sigma(v)$ can be positive, because of Lemma 1. Consequently, we have

$$-1 < \lambda < 1.$$

Let $z_t = \frac{1}{2}(x_t + y_t)$; then $z_t = x_0 + \xi t w$, where $\|w\| = 1$ and $\xi = \frac{1}{2}(1+\lambda)^{1/2}$. Thus

$$0 < \xi < 1.$$

We thus have a point p_w on the sphere Σ defined by $p_w = x_0 + w$. Now, $r(z_t) \geqq \frac{1}{2}[r(x_t) + r(y_t)]$, by Theorem 1. Hence

$$s(z_t) = \frac{r(z_t)}{\xi t} \geqq \frac{1}{2\xi}\left[\frac{r(x_t)}{t} + \frac{r(y_t)}{t}\right] = \frac{1}{2\xi}[s(x_t) + s(y_t)],$$

and

$$\lim_{t \to 0} s(z_t) \geqq \frac{1}{2\xi} \lim_{t \to 0} [s(x_t) + s(y_t)],$$

from which

$$\sigma(w) \geqq \frac{1}{2\xi}[\sigma(u) + \sigma(v)] > \frac{1}{2}[\sigma(u) + \sigma(v)].$$

Thus the theorem is established.

Let $\bar{\sigma}$ denote the least upper bound of the function $\sigma(u)$ as p_u varies over the sphere Σ. Then, also $0 \leqq \bar{\sigma} \leqq 1$; and $\bar{\sigma} = 0$ is possible only for sets which do not have any inner points. For a convex body $\mathfrak{K}$, we have $0 < \bar{\sigma} \leqq 1$.

In the material which follows, it is to be understood that x_0 is a *fixed* boundary point of the convex body $\mathfrak{K}$, $s(x)$ is defined relative to x_0, Σ is the unit sphere about x_0, $\sigma(u)$ is the function defined above on the boundary of Σ, and $\bar{\sigma}$ the least upper bound of $\sigma(u)$ on Σ.

THEOREM 4. *If there is a point p_u on Σ for which $\sigma(u) = \bar{\sigma}$, this point is unique.*

PROOF. Suppose, if possible, that there were a second point p_v for which $\sigma(v) = \bar{\sigma}$. Then, by Theorem 3, since $\bar{\sigma} > 0$, there would be a point p_w for which

$$\sigma(w) > \tfrac{1}{2}[\sigma(u) + \sigma(v)] = \bar{\sigma}.$$

But since no $\sigma(w)$ can exceed $\bar{\sigma}$, there cannot be a second point p_v of the type described.

THEOREM 5. *Let p_u be a point on Σ for which $\sigma(u)=\bar{\sigma}$. If v satisfies the conditions $\|v\|=1$ and $((u, v))<0$, then the points $z_t=x_0+tv$, $t>0$, are exterior points of $\mathfrak{K}$.*

PROOF. Let $p_u=x_0+u$, $\|u\|=1$, and $p_v=x_0+v$, $\|v\|=1$; and let $((u, v))=-\lambda$, where $\lambda>0$. Assume, if possible, that there is a point $z=x_0+dv$, $d>0$, belonging to $\mathfrak{K}$. Let w be the projection (defined in (I)) of z on the line through x_0 and p_u. Then

$$w = p_u + c(x_0 - p_u),$$

where

$$c = \frac{((z - p_u, x_0 - p_u))}{\|p_u - x_0\|^2} = ((z - x_0 - u, -u))$$
$$= ((dv - u, -u)) = 1 + \lambda d.$$

Hence,

$$w = p_u - (1 + \lambda d)u = x_0 - \lambda d u.$$

On the segment joining x_0 to p_u, let $x_t=x_0+tu$ be an inner point of $\mathfrak{K}$. Let y_t be the projection of x_0 on the line through x_t and z. Then

$$y_t = x_t + k(z - x_t)$$

where

$$k = \frac{((x_0 - x_t, z - x_t))}{\|z - x_t\|^2} = \frac{((-tu, dv - tu))}{\|z - x_t\|^2} = \frac{\lambda t d + t^2}{d^2 + 2\lambda t d + t^2},$$

since $z-x_t=z-x_0+x_0-x_t=dv-tu$ and

$$\|z - x_t\|^2 = d^2 - 2td((u, v)) + t^2 = d^2 + 2\lambda t d + t^2.$$

From the above value of k, it is easily seen that $0<k<1$, which means that y_t is a point of $\mathfrak{K}$. The following are easily established:

$$\|y_t - x_0\|^2 = \frac{t^2 d^2(1 - \lambda^2)}{d^2 + 2\lambda t d + t^2} \neq 0,$$

since $\lambda \neq \pm 1$, and

$$\|z - w\|^2 = d^2(1 - \lambda^2).$$

From these, and previous results, we obtain

$$\frac{\|x_t - x_0\|^2}{\|y_t - x_0\|^2} = \frac{t^2(d^2 + 2\lambda td + t^2)}{t^2d^2(1 - \lambda^2)} = \frac{d^2 + 2\lambda td + t^2}{d^2(1 - \lambda^2)} = \frac{\|z - x_t\|^2}{\|z - w\|^2}.$$

Therefore, we have

$$\frac{\|x_t - x_0\|}{\|y_t - x_0\|} = \frac{\|z - x_t\|}{\|z - w\|}. \tag{3}$$

Now $s(y_t) = r(y_t)/\|y_t - x_0\|$ and $s(x_t) = r(x_t)/\|x_t - x_0\|$, where $r(y_t)$ and $r(x_t)$ denote the radii at the points y_t and x_t, respectively. Also $r(y_t) \geqq (1-k)r(x_t)$, by Theorem 1 and the definition of y_t. Hence

$$\begin{aligned} \frac{s(y_t)}{s(x_t)} = \frac{r(y_t)}{\|y_t - x_0\|} \cdot \frac{\|x_t - x_0\|}{r(x_t)} &\geqq (1 - k)\frac{\|x_t - x_0\|}{\|y_t - x_0\|} \\ &= (1 - k)\frac{\|z - x_t\|}{\|z - w\|}, \end{aligned} \tag{4}$$

the last equality being a consequence of (3).

But $k = \|y_t - x_t\|/\|z - x_t\|$ and $1 - k = \|z - y_t\|/\|z - x_t\|$. Therefore, from (4), we have

$$s(y_t) \geqq \frac{\|z - y_t\|}{\|z - w\|} s(x_t). \tag{5}$$

Now,

$$\lim_{t \to 0} \frac{\|z - y_t\|}{\|z - w\|} = \frac{d}{d(1 - \lambda^2)^{1/2}} = \frac{1}{(1 - \lambda^2)^{1/2}} > 1,$$

since z and w are independent of t, while $y_t \to x_0$ as $t \to 0$. Therefore, from (5),

$$\lim_{t \to 0} s(y_t) \geqq \frac{1}{(1 - \lambda^2)^{1/2}} \sigma(u) > \sigma(u) = \bar{\sigma}.$$

But this is impossible; hence the assumption that z was a point of $\mathfrak{K}$ is untenable.

THEOREM 6. *Let p_u be a point on Σ for which $\sigma(u) = \bar{\sigma}$. Then the plane*

$$\pi(x) \equiv ((u, x - x_0)) = 0 \tag{6}$$

is a supporting plane of $\mathfrak{K}$ through the boundary point x_0.

PROOF. If the plane (6) were not a supporting plane, there would be

a point z of $\mathfrak{K}$ for which $\pi(z)<0$. Let $v=(z-x_0)/\|z-x_0\|$; then

$$((u, v)) = \frac{\pi(z)}{\|z - x_0\|} < 0, \qquad z = x_0 + \|z - x_0\|v.$$

But, v satisfies the conditions of Theorem 5; therefore z must be an exterior point of $\mathfrak{K}$. Consequently, there cannot be a point z of $\mathfrak{K}$ for which $\pi(z)<0$; and (6) is indeed a supporting plane.

THEOREM 7. *Let x_0 be a given boundary point of the convex body $\mathfrak{K}$, and let Σ be the unit sphere about x_0. There is a unique point $p_{\bar{u}}$ on Σ for which $\sigma(\bar{u})=\bar{\sigma}$.*

PROOF. We have only to show the existence of one point $p_{\bar{u}}$ for which $\sigma(\bar{u})=\bar{\sigma}$. The uniqueness of this point will follow from Theorem 4.

From the definition of $\bar{\sigma}$ it follows that for any preassigned $\epsilon>0$, there exists a point on Σ for which the value of σ is greater than $\bar{\sigma}-\epsilon$. Choose a monotone decreasing sequence of positive numbers $\{\epsilon_n\}$ with limit zero. Corresponding to each ϵ_n there exists a point p_{u_n} on Σ for which $\sigma(u_n)>\bar{\sigma}-\epsilon_n$. We wish to show that the sequence of points $\{p_{u_n}\}$ on Σ converges.

Let $p_{u_n}=x_0+u_n$, $\|u_n\|=1$, and $p_{u_m}=x_0+u_m$, $\|u_m\|=1$. Then

$$\|p_{u_n} - p_{u_m}\|^2 = 2 - 2((u_n, u_m)). \tag{7}$$

Let $w=\frac{1}{2}(1/\xi)(u_n+u_m)$, where ξ is so chosen that $\|w\|=1$. Then we have

$$\xi^2 = \tfrac{1}{2}[1 + ((u_n, u_m))]. \tag{8}$$

Let $p_w=x_0+w$; from the proof of Theorem 3, we know that

$$\sigma(w) \geqq \frac{1}{2\xi}[\sigma(u_n) + \sigma(u_m)] > \frac{1}{2\xi}[2\bar{\sigma} - \epsilon_n - \epsilon_m].$$

But $\bar{\sigma}\geqq\sigma(w)$; hence $\bar{\sigma}>(1/\xi)[\bar{\sigma}-(\epsilon_n+\epsilon_m)/2]$, from which

$$\xi^2 > \left[1 - \frac{1}{2\bar{\sigma}}(\epsilon_n + \epsilon_m)\right]^2.$$

Using the value of ξ^2 from (8) we easily find that

$$((u_n, u_m)) > 2\left[1 - \frac{1}{2\bar{\sigma}}(\epsilon_n + \epsilon_m)\right]^2 - 1.$$

Then using (7), we obtain

$$\|p_{u_n} - p_{u_m}\|^2 < \frac{4}{\bar{\sigma}}(\epsilon_n + \epsilon_m) - \frac{1}{\bar{\sigma}^2}(\epsilon_n + \epsilon_m)^2. \tag{9}$$

Since $\lim_{n,m\to\infty} \|p_{u_n} - p_{u_m}\| = 0$ and the space $\mathfrak{S}$ is complete, as was assumed in (I) and throughout this paper, the sequence $\{p_{u_n}\}$ converges to a point $p_{\bar{u}}$. This point $p_{\bar{u}}$ is on Σ, and moreover $\sigma(\bar{u}) = \bar{\sigma}$, since it is easily shown that $\sigma(\bar{u})$ is greater than $\bar{\sigma} - \epsilon$ for any preassigned positive ϵ.

THEOREM 8. *A convex body $\mathfrak{K}$ is completely supported at its boundary points.*

PROOF. Let x_0 be a boundary point of $\mathfrak{K}$. There exists a point p_u on the unit sphere Σ about x_0 for which $\sigma(u) = \bar{\sigma}$, by Theorem 7. Hence the plane $((u, x - x_0)) = 0$ is a supporting plane of $\mathfrak{K}$ through x_0, by Theorem 6. Since similar statements can be made for each boundary point, $\mathfrak{K}$ is completely supported at its boundary points.

From the material above, the following additional result may be established without much difficulty:

COROLLARY 3. *There exists a unique supporting plane through the boundary point x_0 of the convex body $\mathfrak{K}$ only if $\bar{\sigma} = 1$; for $\bar{\sigma} < 1$, there is an infinite number of supporting planes through x_0.*

A primary classification of boundary points of a convex body may thus be made in terms of $\bar{\sigma}$, which is a function defined over the boundary of the convex body.

CARNEGIE INSTITUTE OF TECHNOLOGY

A GENERALIZATION OF OSTROWSKI'S THEOREM ON MATRIC IDENTITIES[1]

NEAL H. McCOY

The purpose of this note is to generalize a recent theorem due to Ostrowski[2] which is itself a generalization of a theorem proved by Phillips in 1919.[3] We shall first indicate the nature of Ostrowski's result.

Let $A_1=I, A_2, \cdots, A_m$ be square matrices of order n, I being the unit matrix, and let $x_1, \cdots, x_m$ be numerical parameters. Denote by $F(x_1, \cdots, x_m)$ the determinant of the matrix

$$x_1A_1 + x_2A_2 + \cdots + x_mA_m. \tag{1}$$

Let $\Phi(x_1, \cdots, x_m)$ be the greatest common divisor of the n^2 minors of order $n-1$ of the matrix (1), and set $F/\Phi=F^*(x_1, \cdots, x_m)$. We may now state the theorem of Ostrowski in the following form:[4]

THEOREM 1. *If $B_1, \cdots, B_m$ are matrices of order n, commutative with each other and satisfying the equation*

$$A_1B_1 + A_2B_2 + \cdots + A_mB_m = 0, \tag{2}$$

then

$$F^*(B_1, \cdots, B_m) = 0.$$

Further, if $\Psi(x_1, \cdots, x_m)$ is any polynomial with the property that $\Psi(B_1, \cdots, B_m)=0$ for every set of commutative matrices satisfying (2), *then $\Psi(x_1, \cdots, x_m)$ is divisible by $F^*(x_1, \cdots, x_m)$.*

In this theorem it is tacitly assumed that the elements of the matrices as well as the coefficients of the polynomials are real or complex numbers. In Theorem 3 below we find an extension of the first part of Theorem 1, valid if the elements and coefficients are in an arbitrary commutative ring R with unit element 1. To generalize the second part of the theorem, we find it necessary to make an additional restriction on R, namely, that there exists no nonzero polynomial $\phi(\lambda)$,

[1] Presented to the Society, September 8, 1939.

[2] A. Ostrowski, *On a theorem concerning identical relations between matrices*, Quarterly Journal of Mathematics, vol. 9 (1938), pp. 241–245.

[3] H. B. Phillips, *Functions of matrices*, American Journal of Mathematics, vol. 41 (1919), pp. 266–278.

[4] The assumption that $A_1=I$ is not strictly necessary but assures us that $F(x_1, \cdots, x_m)$ does not vanish identically. For the generalization below, we wish to have $A_1=I$ and so we state the theorem at once in this form.

with coefficients in R, such that $\phi(a)=0$ for all elements a of R. The result obtained under this restriction is stated as Theorem 4.

The form of our theorems is suggested by a recent generalization, in another direction, of Frobenius' theorem concerning the minimum equation of a matrix.[5] Since this plays an important part in the proof of Theorem 4, we state it explicitly before proceeding.

THEOREM 2. *Let R be an arbitrary commutative ring with unit element 1, and A a matrix of order n with elements in R. Let λ be an indeterminate, denote by $f(\lambda)$ the determinant of the matrix $\lambda I-A$, and let $h_{ij}(\lambda)$ be the minors of $\lambda I-A$ of order $n-1$. Then, an element $g(\lambda)$ of $R[\lambda]$ has the property that $g(A)=0$, if and only if $g(\lambda)h_{ij}(\lambda)\equiv 0\ (f(\lambda))$, $(i, j=1, 2, \cdots, n)$.*

It will be seen that Theorems 3 and 4 bear roughly the same relation to Ostrowski's theorem that Theorem 2 does to Frobenius'.

Henceforth we shall let $A_1=I, A_2, \cdots, A_m$ be matrices of order n with elements in a commutative ring R with unit element 1, and let $x_1, \cdots, x_m$ be indeterminates. Denote by $F(x_1, \cdots, x_m)$ the determinant of the matrix

$$x_1A_1 + x_2A_2 + \cdots + x_mA_m. \tag{3}$$

Let $F_{ij}(x_1, \cdots, x_m)$ be the elements of the adjoint of this matrix, and denote by $\mathfrak{m}$ the ideal of those elements $f(x_1, \cdots, x_m)$ of the ring $R[x_1, \cdots, x_m]$ such that

$$f(x_1, \cdots, x_m)F_{ij}(x_1, \cdots, x_m) \equiv 0\ (F(x_1, \cdots, x_m)),\ i, j = 1, 2, \cdots, n.$$

We may now state the following theorem:

THEOREM 3. *If $f(x_1, \cdots, x_m)\equiv 0\ (\mathfrak{m})$, and $B_1, \cdots, B_m$ are commutative matrices of order n, with elements in R, such that*

$$A_1B_1 + A_2B_2 + \cdots + A_mB_m = 0, \tag{4}$$

then $f(B_1, \cdots, B_m)=0$.

The proof is a simple modification of Ostrowski's, and will be only briefly indicated, using his notation so far as possible. Set $A_\mu=(a_{ik}^{(\mu)})$. Now, by hypothesis, we have equations of the form

$$f(x_1, \cdots, x_m)F_{ij}(x_1, \cdots, x_m) = h_{ij}(x_1, \cdots, x_m)F(x_1, \cdots, x_m). \tag{5}$$

But

[5] N. H. McCoy, *Concerning matrices with elements in a commutative ring*, this Bulletin, vol. 45 (1939), pp. 280–284.

$$\sum_{j=1}^{n}\left(\sum_{\mu=1}^{m} x_\mu a_{jk}^{(\mu)}\right)F_{ji}(x_1, \cdots, x_m) = \delta_i^k F(x_1, \cdots, x_m).$$

Multiply this last equation by $f(x_1, \cdots, x_m)$ and use (5), getting

$$F(x_1, \cdots, x_m)\sum_{j=1}^{n}\left(\sum_{\mu=1}^{m} x_\mu a_{jk}^{(\mu)}\right)h_{ji}(x_1, \cdots, x_m) = \delta_i^k F(x_1, \cdots, x_m) f(x_1, \cdots, x_m). \tag{6}$$

Now, if $F(x_1, \cdots, x_m)$ is arranged in terms of decreasing powers of x_1, it is clear that the first term is x_1^n. It follows readily that, in the ring $R[x_1, \cdots, x_m]$, $F(x_1, \cdots, x_m)$ is not a divisor of zero, and can therefore by divided out of equation (6), yielding

$$\sum_{j=1}^{n}\left(\sum_{\mu=1}^{m} x_\mu a_{jk}^{(\mu)}\right)h_{ji}(x_1, \cdots, x_m) = \delta_i^k f(x_1, \cdots, x_m).$$

Let $C_{jk} = \sum_{\mu=1}^{m} B_\mu a_{jk}^{(\mu)}$. Then we have from the preceding equation, and the fact that the B's are commutative,

$$\sum_{j=1}^{n} C_{jk} h_{ji}(B_1, \cdots, B_m) = \delta_i^k f(B_1, \cdots, B_m).$$

Now this equation corresponds to Ostrowski's equation (2.3), and the remainder of the proof will be omitted as from this point the proof coincides with his.

If R is a field, or more generally, a domain of integrity with unique factorization into primes, then it follows readily that $\mathfrak{m}$ is the principal ideal $(F(x_1, \cdots, x_m)/D(x_1, \cdots, x_m))$, where $D(x_1, \cdots, x_m)$ is the greatest common divisor of the $F_{ij}(x_1, \cdots, x_m)$. In this case, our Theorem 3 can be expressed in the same form as the first part of Theorem 1.

Before proceeding, we pause to make a remark which will indicate how, in another way, the ideal $\mathfrak{m}$ has properties generalizing the familiar properties of the minimum function of a single matrix with elements in a field. From the definition of $\mathfrak{m}$, it follows that if $g(x_1, \cdots, x_m) \equiv 0\ (\mathfrak{m})$, then

$$g(x_1, \cdots, x_m)\ \text{adj}\ (x_1A_1 + \cdots + x_mA_m) = KF(x_1, \cdots, x_m),$$

where K is a matrix with elements in $R[x_1, \cdots, x_m]$. By taking determinants and dividing by F^{n-1}, we see that

$$[g(x_1, \cdots, x_m)]^n \equiv 0\ (F(x_1, \cdots, x_m)).$$

In particular, it follows at once that in $R[x_1, \cdots, x_m]$ the ideals $\mathfrak{m}$

and $(F(x_1, \cdots, x_m))$ have precisely the same prime ideal divisors.

We now turn to the problem of generalizing the second part of Theorem 1, and we find it necessary to make the following additional assumption concerning R. *If $f(\lambda)$ is an element of $R[\lambda]$ such that for every element a of R, $f(a)=0$, then $f(\lambda)=0$.* It follows readily by induction on the number of indeterminates that if in $R[x_1, \cdots, x_k]$, $f(x_1, \cdots, x_k)$ has the property that $f(a_1, \cdots, a_k)=0$ for all choices of $a_1, \cdots, a_k$ in R, then $f(x_1, \cdots, x_k)=0$, $(k=1, 2, \cdots)$. We now prove the following lemma:

LEMMA. *Let $x_1, \cdots, x_m$ be indeterminates, and consider an element of $R[x_1, \cdots, x_m]$ of the form*

$$g(x_1, \cdots, x_m) = x_1^n + g_1(x_2, \cdots, x_m)x_1^{n-1} + \cdots + g_n(x_2, \cdots, x_m).$$

If $f(x_1, \cdots, x_m)$ is an element of $R[x_1, \cdots, x_m]$ such that for every choice of $a_2, \cdots, a_m$ in R, $f(x_1, a_2, \cdots, a_m)$ is divisible by $g(x_1, a_2, \cdots, a_m)$ in $R[x_1]$, then $f(x_1, \cdots, x_m)$ is divisible by $g(x_1, \cdots, x_m)$ in the ring $R[x_1, \cdots, x_m]$.

Let x represent the composite set $x_2, \cdots, x_m$; and a the set of elements $a_2, \cdots, a_m$ of R. Then we may write

$$g(x_1, \cdots, x_m) = x_1^n + g_1(x)x_1^{n-1} + \cdots + g_n(x),$$

and

$$f(x_1, \cdots, x_m) = f_0(x)x_1^p + f_1(x)x_1^{p-1} + \cdots + f_p(x).$$

Now, by hypothesis, we have for arbitrary but fixed a a relation of the form

$$\begin{aligned} f_0(a)x_1^p + \cdots + f_p(a) = {} & [x_1^n + g_1(a)x_1^{n-1} + \cdots + g_n(a)] \\ & \cdot [h_0x_1^{p-n} + h_1x_1^{p-n-1} + \cdots + h_{p-n}], \end{aligned} \tag{7}$$

where the h_i are elements of R. But equation (7) is equivalent to a set of $p+1$ equations in R, obtained by equating the coefficients of the different powers of x_1 on each side of (7). These equations take the form

$$\begin{aligned} f_0(a) &= h_0, \\ f_1(a) &= h_1 + h_0g_1(a), \\ &\cdots\cdots\cdots\cdots, \\ f_{p-n}(a) &= h_{p-n} + h_{p-n-1}g_1(a) + \cdots + h_0g_{p-n}(a), \end{aligned} \tag{8}$$

together with the set of equations

$$f_t(a) = \sum_{i+j=t} h_i g_j(a), \qquad t = p - n + 1, \cdots, p. \tag{9}$$

The equations (8) can be solved in turn for the h_i, and these unique solutions take the form

$$h_i = G_i[f_1(a), \cdots, f_{p-n}(a); g_1(a), \cdots, g_{p-n}(a)], \quad i = 0, 1, \cdots, p - n,$$

where the G_i are polynomials with integral coefficients. Thus, in (7), the coefficients h_i are uniquely determined by the choice of a. Now let us set

$$h_i(x) = G_i[f_1(x), \cdots, f_{p-n}(x); g_1(x), \cdots, g_{p-n}(x)], \quad i = 0, 1, \cdots, p - n, \tag{10}$$

so that our original h_i is $h_i(a)$. Then equations (9) state that for *every* a in R,

$$f_t(a) - \sum_{i+j=t} h_i(a) g_j(a) = 0,$$

and therefore, by hypothesis on R, it follows that

$$f_t(x) = \sum_{i+j=t} h_i(x) g_j(x), \qquad t = p - n + 1, \cdots, p. \tag{11}$$

But equations (10) and (11) are precisely the set of equations which state that

$$f(x_1, \cdots, x_m) = g(x_1, \cdots, x_m)[h_0(x)x_1^{p-n} + \cdots + h_{p-n}(x)],$$

and the lemma is established.

We shall now prove the following theorem under the assumption on R which we have made throughout this section:

THEOREM 4. *If $\phi(x_1, \cdots, x_m)$ is an element of $R[x_1, \cdots, x_m]$ with the property that $\phi(B_1, \cdots, B_m) = 0$ for every $B_1, \cdots, B_m$ which are commutative and satisfy* (4), *then $\phi(x_1, \cdots, x_m) \equiv 0$* (m).

In the proof of this theorem we shall not distinguish between the ring R and the ring of matrices of the form aI, where a is in R. Accordingly, we identify I with 1, the unit element of R.

Let $a_2, \cdots, a_m$ be arbitrary elements of R, and let us choose

$$B_1 = -a_2A_2 - \cdots - a_mA_m, \quad B_2 = a_2, \cdots, \quad B_m = a_m. \tag{12}$$

Then clearly condition (4) is satisfied, and by hypothesis we have

$$\phi(B_1, B_2, \cdots, B_m) = \phi(B_1, a_2, \cdots, a_m) = 0.$$

Now if $F(x_1, \cdots, x_m)$ and $F_{ij}(x_1, \cdots, x_m)$ have the same meaning as above, then clearly the determinant of $x_1 - B_1$ is $F(x_1, a_2, \cdots, a_m)$, and the first minors of the matrix $x_1 - B_1$ are precisely the $F_{ij}(x_1, a_2, \cdots, a_m)$, except possibly for sign.

Since $\phi(B_1, a_2, \cdots, a_m) = 0$, it follows by Theorem 2 that, in the ring $R[x_1]$, $\phi(x_1, a_2, \cdots, a_m)F_{ij}(x_1, a_2, \cdots, a_m)$ is divisible by $F(x_1, a_2, \cdots, a_m)$, $(i, j = 1, 2, \cdots, n)$. But $F(x_1, \cdots, x_m)$ is of the form of the $g(x_1, \cdots, x_m)$ in the lemma, and hence the lemma can be applied as the above holds for all choices of $a_2, \cdots, a_m$ in R. Thus

$$\phi(x_1, \cdots, x_m)F_{ij}(x_1, \cdots, x_m) \equiv 0\,(F(x_1, \cdots, x_m)),\ i, j = 1, 2, \cdots, n,$$

that is,

$$\phi(x_1, \cdots, x_m) \equiv 0\ (\mathfrak{m}).$$

This completes the proof of the theorem.

We have established Theorem 4, following Ostrowski, not in fact under the assumption that $\phi(B_1, \cdots, B_m) = 0$ for *every* $B_1, \cdots, B_m$ which are commutative and satisfy (4), but under the weaker assumption that the B's may be restricted to be in the special form (12). Because of the homogeneity of the polynomials considered, it is not difficult to show that we can further restrict our hypothesis by assuming always that $a_m = 1$. If, in the lemma, we consider only homogeneous polynomials, we can also in it assume that $a_m = 1$. Thus, for the case $m = 2$, the lemma and Theorem 4 are true for *any* commutative ring R with unit element. Theorem 4, so interpreted, then yields an actual generalization of Theorem 2.

Note added in proof: The assumption that $A_1 = I$ is used, so far as Theorem 3 is concerned, only to make sure that $F(x_1, \cdots, x_m)$ is not a divisor of zero in $R[x_1, \cdots, x_m]$. The remarks to follow will show that it is certainly sufficient, although by no means necessary, to assume that for some i, $|A_i|$ is not a divisor of zero in R.

If $ab = 0$, $a \neq 0$, a is called an *annihilator* of b. It is quite easy to prove, although this fact does not seem to be in the literature, that *if* $g(x_1, \cdots, x_m)$ *is an element of* $R[x_1, \cdots, x_m]$ *with an annihilator* $h(x_1, \cdots, x_m)$ *in* $R[x_1, \cdots, x_m]$, *then* $g(x_1, \cdots, x_m)$ *has an annihilator* a *in* R. Otherwise expressed, $g(x_1, \cdots, x_m)$ is a divisor of zero in $R[x_1, \cdots, x_m]$ if and only if all coefficients in g are annihilated by the *same* element a of R.

SMITH COLLEGE

THE RADIUS AND MODULUS OF n-VALENCE FOR ANALYTIC FUNCTIONS WHOSE FIRST $n-1$ DERIVATIVES VANISH AT A POINT

LYNN H. LOOMIS

The principal result of this note is the determination of the precise radius and modulus of n-valence for the class of functions $f(z)=z^n+a_{n+1}z^{n+1}+\cdots$ analytic and less than or equal to M in modulus in $|z|\leqq 1$. This result readily leads to the radius and modulus of n-valence for the more general class of functions $f(z)=az^n+a_{n+1}z^{n+1}+\cdots$ analytic and less than or equal to M in modulus in $|z|\leqq R$. Finally, we note certain approximations which rather naturally suggest themselves in a search for more easily calculable constants.

We consider only expansions about the origin of functions $f(z)$ with $f(0)=0$, the generalization to expansions about a of functions $f(z)$ with $f(a)=b$ being obvious. Each circle mentioned will be understood to have the origin ($w=0$ or $z=0$) as center. The phrases *radius of n-valence* and *modulus of n-valence*, which usually refer to a class of functions, will also be used with reference to a single function. The radius of n-valence of the function $f(z)$ is the radius of the largest circle within which $f(z)$ assumes no value more than n times, and assumes at least one value n times. The modulus of n-valence of $f(z)$ is the radius of the largest circle of which the interior is covered exactly n times by the map under $f(z)$ of $|z|<\rho$, where ρ is the above radius of n-valence. Consider now one of the classes defined above. It is obvious that for each function $w=f(z)$ of the class there is a neighborhood of $z=0$ in which the function assumes no value more than n times, and assumes exactly n times every value in a sufficiently small neighborhood of $w=0$. The radius of n-valence ρ_n of the class is the radius of the largest circle within which *no* function of the class assumes a value more than n times. The modulus of n-valence m_n of the class is the radius of the largest circle of which the interior is covered exactly n times by the map of $|z|<\rho_n$ under *every* function of the class.

THEOREM. *Consider the class of functions $f(z)=z^n+a_{n+1}z^{n+1}+\cdots$ analytic and less than or equal to M $(M>1)$[1] in modulus in $|z|\leqq 1$,*

[1] The restriction to $M>1$ is necessary. By the Cauchy coefficient inequality, $M\geqq 1$, and if $M=1$ the class consists of the single function $f(z)=z^n$ for which the theorem is false.

where n and M are the constants of the class. The radius ρ_n and the modulus m_n of n-valence of the class are given by

$$m_n = \frac{M\rho_n^n(1 - M\rho_n)}{M - \rho_n}, \qquad \rho_n = M_n - (M_n^2 - 1)^{1/2},$$

where

$$M_n = \frac{1}{2}\left[\left(1 + \frac{1}{n}\right)M + \left(1 - \frac{1}{n}\right)\frac{1}{M}\right].$$

The case for $n=1$ is completely treated in the literature. The results are due to Landau and Dieudonné.[2] The values of ρ_1 and m_1 are

$$\rho_1 = M - (M^2 - 1)^{1/2}, \qquad m_1 = M\rho_1^2 = M\rho_1\frac{(1 - M\rho_1)}{M - \rho_1}.$$

The following proof for general n consists of

(a) a proof that every function of the class is n-valent and covers $|w| < m_n$ exactly n times in $|z| < \rho_n$, and

(b) the exhibition of a single function of the class for which ρ_n and m_n are respectively the radius and modulus of n-valence.

We employ the device which Dieudonné used to find the radius of star-shapedness in the case $n=1$ (Montel, loc. cit., p. 94), and then apply a theorem due to S. Ozaki,[3] which states that if $f(z)$ is analytic in $|z| \leqq r$ and has n zeros there, none on the circumference, and if for some real α, $\Re[e^{i\alpha}zf'(z)/f(z)] > 0$ on $|z| = r$, then $f(z)$ is n-valent in the circle $|z| < r$.

Consider

$$g(z) = f(z)/z^n = 1 + a_{n+1}z + \cdots.$$

Since $g(z)$ is analytic and less than or equal to M in modulus in $|z| \leqq 1$ with $g(0) = 1$, the following inequalities, results of the Schwarz lemma, are valid (Montel, loc. cit., p. 91):

$$(1) \qquad |g(z)| \geqq \frac{M(1 - Mr)}{M - r},$$

$$(2) \qquad |g'(z)| \leq \frac{M^2 - |g(z)|^2}{M(1 - r^2)},$$

[2] See Montel, *Leçons sur les Fonctions Univalentes ou Multivalentes*, 1933, pp. 90–95. This book contains a convenient collection of material relating to this paper, and we shall often refer to it.

[3] S. Ozaki, *Some remarks on the univalency and multivalency of functions*, Science Reports of Tokyo Bunrika Daigaku, section A, 2, no. 32, 1934, pp. 41–55.

where $r=|z|<1/M$. From (1) and (2),

$$\left|\frac{zg'(z)}{g(z)}\right| \leqq \frac{r(M^2-1)}{(M-r)(1-Mr)} \tag{3}$$

if $r<1/M$. But we have

$$g'(z)=\frac{z^n f'(z)-nz^{n-1}f(z)}{z^{2n}},$$

$$\frac{zf'(z)}{f(z)}=\frac{zg'(z)}{g(z)}+n. \tag{4}$$

By (3) and (4) $\Re[zf'(z)/f(z)]>0$ whenever

$$n>\frac{r(M^2-1)}{(M-r)(1-Mr)}.$$

The latter inequality reduces, for $r<1/M$, to

$$r^2-\left[\left(1+\frac{1}{n}\right)M+\left(1-\frac{1}{n}\right)\frac{1}{M}\right]r+1>0,$$

which is satisfied (factoring the left member) if

$$r<M_n-(M_n^2-1)^{1/2}=\rho_n,$$

where

$$M_n=\frac{1}{2}\left[\left(1+\frac{1}{n}\right)M+\left(1-\frac{1}{n}\right)\frac{1}{M}\right].$$

Now $\rho_n<1/M$, for we shall see that ρ_n is the smallest positive zero of the derivative of

$$Mx^n(1-Mx)/(M-x),$$

and this lies between 0 and $1/M$. But from (1)

$$|f(z)| \geqq \frac{Mr^n(1-Mr)}{M-r}. \tag{5}$$

Hence $f(z)$ has precisely n zeros in $|z|<r$ and none on the boundary $|z|=r$, for every $r\leqq\rho_n$. Then:

(1) By (5), $|f(z)|\geqq m_n$ on $|z|=\rho_n$. Thus, by Rouché's theorem, $f(z)$ assumes exactly n times in $|z|<\rho_n$ every value w with $|w|<m_n$.

(2) We have seen that $\Re[zf'(z)/f(z)]>0$ on $|z|=r<\rho_n$.

Thus, by the theorem of Ozaki noted above, $f(z)$ is n-valent in the

region $|z|<r$ for every $r<\rho_n$, and hence n-valent in $|z|<\rho_n$. This completes the proof of (a).

It follows from the Schwarz lemma that equality can occur in (1) for a point in $0<|z|<1$ only if $g(z)$ is of the form

$$g(z)=\frac{M(1-Me^{i\theta}z)}{M-e^{i\theta}z}.$$

We obtain the simplest function with $e^{i\theta}=1$, giving

$$f_0(z)=\frac{Mz^n(1-Mz)}{M-z}.$$

This function is of the class considered, and has ρ_n and m_n for its radius and modulus of n-valence. For,

$$f_0'(z)=\frac{M^2nz^{n-1}}{(M-z)^2}\left\{z^2-\left[\left(1+\frac{1}{n}\right)M+\left(1-\frac{1}{n}\right)\frac{1}{M}\right]z+1\right\},$$

and it is evident that the zero of $f_0'(z)$ nearest the origin (except $z=0$ itself) is precisely ρ_n. Also

$$f_0(\rho_n)=\frac{M\rho_n^n(1-M\rho_n)}{M-\rho_n}=m_n.$$

Thus these constants are respectively the radius and modulus of n-valence for this function, and hence by (a) for the class considered in the theorem.

COROLLARY 1. *Consider the class of functions* $f(z)=az^n+a_{n+1}z^{n+1}+\cdots$ *analytic and less than or equal to* M $(M>|a|R^n)$ *in modulus in* $|z|\leqq R$, *the constants of the class being* $|a|$ $(\neq 0)$, n, M *and* R. *The radius* ρ_n *and the modulus* m_n *of* n*-valence of the class are given by*

$$\rho_n=R\sigma,\qquad m_n=\frac{M\sigma^n(|a|R^n-M\sigma)}{M-|a|R^n\sigma},$$

where σ *is defined by*

$$\sigma=M_n-(M_n^2-1)^{1/2},$$

$$M_n=\frac{1}{2}\left[\left(1+\frac{1}{n}\right)\frac{M}{|a|R^n}+\left(1-\frac{1}{n}\right)\frac{|a|R^n}{M}\right].$$

This corollary follows immediately from the theorem upon considering the function $g(z)=f(Rz)/aR^n$.

The results on the modulus of n-valence relate to a result due to Walsh and Seidel,[4] who do not assume $f'(0)=\cdots=f^{(n-1)}(0)=0$, but who, on the other hand, do not obtain the sharp inequality.

COROLLARY 2. *For the class of functions of Corollary* 1

$$\rho_n > \frac{|a|R^{n+1}}{2M}, \qquad m_n > M\left(\frac{|a|R^n}{2M}\right)^{n+1}.$$

For, $M/|a|R^n>1$, from which it follows that $M/|a|R^n \geqq M_n>1$. Therefore, since

$$M_n-(M_n^2-1)^{1/2}=\frac{1}{M_n+(M_n^2-1)^{1/2}}>\frac{1}{2M_n},$$

we find that

$$\frac{|a|R^{n+1}}{2M}<\rho_n. \tag{6}$$

Moreover, if $f(z)=Mz^n(|a|R^{n+1}-Mz)/(MR^{n+1}-|a|R^{2n}z)$, then $f(\rho_n)=m_n$, $f'(\rho_n)=0$, $f(|a|R^{n+1}/2M)<m_n$ by (6), and the last inequality easily gives

$$M\left(\frac{|a|R^n}{2M}\right)^{n+1}<m_n. \tag{7}$$

But (6) and (7) constitute the corollary.

The inequality (7) is an improvement over an approximation due to Privaloff,[5] which states that the image of the circle $|z|<R$ under the function $w=f(z)$ of the class of Corollary 1 covers at least n times the circle

$$|w|<\frac{8}{3}M\left(\frac{R^n|a|}{4M}\right)^{n+1}.$$

For,

$$M\left(\frac{|a|R^n}{2M}\right)^{n+1}>\frac{8}{3}M\left(\frac{|a|R^n}{4M}\right)^{n+1},$$

and by (7), the circle $|w|<M(|a|R^n/2M)^{n+1}$ is covered exactly n times by the image of $|z|<\rho_n$ and hence at least n times by that of $|z|<R$.

[4] Walsh and Seidel, *On the derivatives of functions analytic in the unit circle*, Proceedings of the National Academy of Sciences, vol. 24 (1938), pp. 337–340.

[5] J. Privaloff, *Sur un théorème de M. Bloch*, Recueil Mathématique de Moscou, vol. 35 (1928), pp. 111–121.

We conclude with a few miscellaneous remarks.

1. The modulus of n-valence for the class of the theorem is also the radius of the largest circle within which converge all the power series (in $w^{1/n}$) of the inverse functions.

2. It is evident from the proof of part (b) of the theorem that the radius and modulus of n-valence for the function $f(z)$ are ρ_n and m_n only if

$$f(z) = \frac{Mz^n(1 - Me^{i\theta}z)}{M - e^{i\theta}z}$$

for some real θ. For any other function of the class the radius and modulus of n-valence are greater, respectively, than ρ_n and m_n.

3. The inequalities $\leqq M$, $\leqq 1$ of the theorem may be replaced by $< M$, < 1 without affecting the validity of the work.

4. It is easily seen that the equations of Corollary 1 give, for instance, the radius of n-valence for the class of functions of the form $f(z) = a_f z^n + a_{n+1}z^{n+1} + \cdots$ analytic and less than or equal to M_f in modulus in $|z| \leqq R$, including just those functions for which $M_f/|a_f|$ is less than some preassigned bound ($M/|a|$ of the equations of the corollary). No modulus of n-valence exists for this class.

HARVARD UNIVERSITY

INVERSE PROBLEMS OF THE CALCULUS OF VARIATIONS FOR MULTIPLE INTEGRALS[1]

WILLIAM A. PATTERSON

1. **Introduction.** The simplest case of the inverse problem of Darboux is that in which an ordinary differential equation in the normal form $y''=\phi(x, y, y')$ is assigned with the requirement that we ascertain, first, under what conditions $\phi(x, y, y')$ is the solution for y'' of the Euler equation of a variation problem of the form $\int_{x_1}^{x_2} f(x, y, y')dx = \min$ and then that we determine the most general integrand function f corresponding to an admissible function[2] $\phi(x, y, y')$.

For partial differential equations the simplest analogous problem is that of finding the most general *first order* multiple integral variation problem *associated* with an assigned partial differential equation, that is, the most general integrand function f of a variation problem of the form

$$\text{(I)} \qquad \int_{(R)} f(x_1, \cdots, x_n, z, p_1, \cdots, p_n)dx_1 \cdots dx_n = \min, \quad p_i = \partial z/\partial x_i,$$

of which the extremal hypersurfaces are the integral hypersurfaces $z=z(x_1, \cdots, x_n)$ of a prescribed partial differential equation.

A systematic study of such inverse problems of Darboux type for certain important classes of partial differential equations is made in this paper.

2. **A uniqueness theorem.** Consider a partial differential equation of the form

$$\text{(2.1)} \qquad F \equiv A_{\alpha\beta}(x_1, \cdots, x_n, z, p_1, \cdots, p_n)p_{\alpha\beta} + B(x_1, \cdots, x_n, z, p_1, \cdots, p_n) = 0,$$

where $p_{ij}=\partial^2 z/\partial x_i \partial x_j$, and $A_{ij}=A_{ji}$ $(i, j=1, \cdots, n)$ and B are arbitrary analytic functions of $x_1, \cdots, x_n, z, p_1, \cdots, p_n$. In (2.1) as elsewhere in this paper, a repeated Greek letter is an umbral index indicating a summation with range 1 to n, unless otherwise indicated.

Equation (2.1), as it stands, may have an equation of variation which is self-adjoint on every hypersurface $z=z(x_1, \cdots, x_n)$. If so there is always a multiple integral of the form (I) having $F=0$ as its

[1] Presented to the Society, December 29, 1939.

[2] Cf. G. Darboux, *Théorie des Surfaces*, vol. 3, 1887, p. 53. For the case of $n \geqq 2$ dependent variables $y_1, \cdots, y_n$ see L. LaPaz, Proceedings of the National Academy of Sciences, vol. 17 (1931), pp. 459–463.

Euler-Lagrange equation. The most general such integral is known to have an integrand function of the form[3]

$$f = G + C + D_a p_a + \partial\omega_a/\partial x_a, \tag{2.2}$$

where $G(x_1, \cdots, x_n, z, p_1, \cdots, p_n)$ is a particular solution of

$$g_{p_i p_j} = A_{ij}, \qquad i, j = 1, \cdots, n, \tag{2.3}$$

where the set $\{C(x_1, \cdots, x_n, z), D_i(x_1, \cdots, x_n, z)\}$ is a particular solution of

$$C_z - D_{a z_a} = G_{p_a z_a} + G_{p_a z} p_a - G_z - B, \tag{2.4}$$

and where the ω_i occurring in the partial derivatives with respect to x_i are arbitrary functions of $x_1, \cdots, x_n, z$ alone. Whether (2.1), as it stands, has a self-adjoint equation of variation or not, there may exist equivalent equations which do have self-adjoint equations of variation. We proceed to the question of the existence of such equivalent equations.

For a problem of minimizing the multiple integral (I), the Lagrange partial differential equation must necessarily be a second order partial differential equation[4] linear in the p_{ij} and with a self-adjoint equation of variation. Thus, if the integral hypersurfaces $z = z(x_1, \cdots, x_n)$ of $F = 0$ are to be the extremal hypersurfaces of such a problem, the partial differential equation $F = 0$ must be equivalent to a partial differential equation of the second order with the properties of linearity and self-adjointness of the equation of variation just described. Since the analytic partial differential equation $F = 0$ is itself linear in the p_{ij}, it follows that the most general linear equation

$$a_{\mu\nu}(x_1, \cdots, x_n, z, p_1, \cdots, p_n)p_{\mu\nu} + b(x_1, \cdots, x_n, z, p_1, \cdots, p_n) = 0$$

equivalent to $F = 0$ is of the form $M \cdot F = 0$, where $M \neq 0$ is a function of $x_1, \cdots, x_n, z, p_1, \cdots, p_n$ alone. A function $M(x_1, \cdots, x_n, z, p_1, \cdots, p_n) \neq 0$ and such that the equation $M \cdot F = 0$ has a self-adjoint equation of variation will be called a *multiplier* of the partial differential equation $F = 0$. Since multipliers which differ only by a nonzero constant factor are not regarded as distinct, it is permissible to restrict attention to multipliers $M > 0$.

If the equation $M \cdot F = 0$ is to have a self-adjoint equation of variation, then the multiplier M must satisfy the following relations[5] identically in the variables $x_1, \cdots, x_n, z, p_1, \cdots, p_n$:

[3] Cf. L. LaPaz, Transactions of this Society, vol. 32 (1930), p. 513.

[4] Integrand functions linear in the p_i are excluded from consideration.

[5] L. LaPaz, loc. cit., p. 512.

$$M \cdot A_{ij} = M \cdot A_{ji}, \qquad i, j = 1, 2, \cdots, n, \tag{2.5}$$

$$\begin{aligned} &\partial(M \cdot A_{ij})/\partial p_k = \partial(M \cdot A_{kj})/\partial p_i, \quad i, j, k = 1, 2, \cdots, n;\ k \neq i, \\ &\partial(M \cdot A_{ia})/\partial x_a + p_a \partial(M \cdot A_{ia})/\partial z - \partial(M \cdot B)/\partial p_i = 0, \\ &\qquad\qquad i = 1, 2, \cdots, n. \end{aligned} \tag{2.6}$$

It is evident that the conditions (2.5) are automatically satisfied. Hence if a multiplier $M(x_1, \cdots, x_n, z, p_1, \cdots, p_n)$ exists, it must satisfy the relations (2.6) which may be written in the form

$$\begin{aligned} &A_{ij}\partial M/\partial p_k - A_{kj}\partial M/\partial p_i + (\partial A_{ij}/\partial p_k - \partial A_{kj}/\partial p_i) \cdot M = 0, \\ &A_{ia}\partial M/\partial x_a + p_a \cdot A_{ia}\partial M/\partial z - B \cdot \partial M/\partial p_i \\ &\qquad + (\partial A_{ia}/\partial x_a + p_a \partial A_{ia}/\partial z - \partial B/\partial p_i) \cdot M = 0. \end{aligned} \tag{2.7}$$

If $F=0$, as it stands, has a self-adjoint equation of variation, then (2.7) reduces to

$$\begin{aligned} A_{ij}\partial M/\partial p_k - A_{kj}\partial M/\partial p_i &= 0, \\ A_{ia}\partial M/\partial x_a + p_a \cdot A_{ia}\partial M/\partial z - B\partial M/\partial p_i &= 0. \end{aligned} \tag{2.8}$$

If $\mathfrak{a}$, the determinant of the matrix (A_{ij}), is different from zero, it follows from (2.8_1) that $\partial M/\partial p_i = 0$ $(i=1, 2, \cdots, n)$. But then (2.8_2) is seen to imply that $\partial M/\partial x_i = \partial M/\partial z = 0$ $(i=1, 2, \cdots, n)$; that is to say, M is a constant. This completes the proof of the following result.

THEOREM 2.1. *If $F=0$, as it stands, has a self-adjoint equation of variation in a neighborhood in which $\mathfrak{a} \neq 0$, then $F=0$ admits a constant as its most general multiplier in this neighborhood.*

As an immediate corollary to Theorem 2.1 we have the following important uniqueness theorem.

THEOREM 2.2. *In a neighborhood in which $\mathfrak{a} \neq 0$, the equation $F=0$ admits at most one multiplier M.*

3. **Variation problems in $(n+1)$-space of which the extremals are minimal hypersurfaces.** As an interesting application of Theorem 2.2, let us consider the inverse problem of Darboux for minimal hypersurfaces[6] in $(n+1)$-space $(n>2)$. The corresponding inverse problem for straight lines in $(n+1)$-space $(n>2)$, that is, for solution curves of the system of differential equations $y_i'' = 0$ $(i=1, 2, \cdots, n)$ has recently been treated by D. R. Davis[7] who has found that the most general integrand function $f(x, y_1, \cdots, y_n, y_1', \cdots, y_n')$ of a varia-

[6] For the case $n=2$ see L. LaPaz, Acta Szeged, vol. 5 (1932), pp. 199–207.

[7] D. R. Davis, Transactions of this Society, vol. 33 (1931), pp. 244–251.

tion problem of the form $\int_{x_1}^{x_2} f dx = \min$ for which the extremals are the solution curves of the system of differential equations $y_i'' = 0$ $(i = 1, 2, \cdots, n)$ is of the form $f = g(x, y_1, \cdots, y_n, y_1', \cdots, y_n') + dt(x, y_1, \cdots, y_n)/dx$, where g is a suitably chosen particular solution of a (compatible) system of second order partial differential equations, with $i < j = 1, 2, \cdots, n$, of the form

$$\partial^2 g / \partial y_i' \partial y_j' = P_{ij}(y_1', \cdots, y_n', y_1 - y_1' x, \cdots, y_n - y_n' x),$$

and where t is an arbitrary function of its arguments. The functions P_{ij} are arbitrary solutions of a certain system of partial differential equations considered by Davis. In sharp contrast to this result we shall show that the most general variation problem (I) for which the extremals are minimal hypersurfaces in the sense defined below is uniquely determined up to an additive function of the form $\partial \omega_\alpha / \partial x_\alpha$ where the ω_i $(i = 1, 2, \cdots, n)$ are arbitrary functions of $x_1, \cdots, x_n, z$.

A surface $z = z(x_1, \cdots, x_n)$ defined and of class C'' in a region R of $(x_1, \cdots, x_n)$-space will be called a minimal hypersurface if z and its partial derivatives of the first and second orders satisfy in R the partial differential equation

$$F_1 \equiv [(1 + p_\gamma p_\gamma)\delta_\alpha^\beta - p_\alpha p_\beta] \cdot p_{\alpha\beta} = 0, \tag{3.1}$$

where δ_j^i is Kronecker's delta. $F_1 = 0$ is of the form (2.1) with $a \equiv (1 + p_\gamma p_\gamma)^{n-1} \neq 0$ and therefore by Theorem 2.2 there exists at most one multiplier M for $F_1 = 0$. But the Lagrange partial differential equation for the problem $\int_{(n)} (1 + p_\gamma p_\gamma)^{1/2} dx_1 \cdots dx_n = \min$ is of the form $M \cdot F_1 = 0$ where $M = (1 + p_\gamma p_\gamma)^{-3/2}$. Hence, calculating the integrand function f of (2.2), we reach the following conclusion.

THEOREM 3.1. *The partial differential equation $F_1 = 0$ of minimal hypersurfaces admits the unique multiplier $M = (1 + p_\gamma p_\gamma)^{-3/2}$. The most general variation problem* (I) *for which the extremal hypersurfaces are minimal hypersurfaces has an integrand function f of the form $f = (1 + p_\gamma p_\gamma)^{1/2} + \partial \omega_\alpha / \partial x_\alpha$, where the ω_i are arbitrary functions of $x_1, \cdots, x_n, z$ alone.*

4. **Variation problems associated with the equation $A_{\alpha\beta} p_{\alpha\beta} + B(x_1, \cdots, x_n, z) = 0$.** In this section we shall consider the problem of finding the most general variation problem (I) associated with the partial differential equation

$$F_2 \equiv A_{\alpha\beta} p_{\alpha\beta} + B(x_1, \cdots, x_n, z) = 0, \tag{4.1}$$

where $A_{ij} = A_{ji}$ $(i, j = 1, 2, \cdots, n)$ are real constants not all zero and B is an arbitrary analytic function of $x_1, \cdots, x_n, z$.

Partial differential equations of the form (4.1) are of much importance in pure and applied mathematics. For example if we choose $n=2$, $A_{ij}=\delta_j^i$ and $B=8z/(1+x_1^2+x_2^2)^2$, (4.1) reduces to the partial differential equation studied by Schwarz in his researches on minimal surfaces. For the same value of n and suitable choice of A_{ij} and B, the equation (4.1), with appropriate interpretation of the variables x_1, x_2 and z, includes many of the most important partial differential equations of mathematical physics.

It is easy to verify that all equations (4.1) have self-adjoint equations of variation and hence, as they stand, are Euler-Lagrange equations. We therefore proceed immediately to the problem of determining the most general integrand function f of the problem (I) known to be associated with an equation of the form $F_2=0$. Application of Theorem 2.1 and calculation of the function f of (2.2) leads to the following theorem.

THEOREM 4.1. *If* $\mathfrak{a}\neq 0$, *then the partial differential equation* $F_2=0$ *admits a constant as its most general multiplier. The most general variation problem* (I) *associated with* $F_2=0$ *has an integrand function* f *of the form* $f=\frac{1}{2}A_{\alpha\beta}p_\alpha p_\beta-\int_{z_0}^{z}B\,dz+\partial\omega_\alpha/\partial x_\alpha$, *where the* ω_i *are arbitrary functions of* $x_1, \cdots, x_n, z$.

As an application of this case we find that the most general variation problem (I) associated with $s=\partial^2 z/\partial x_1\partial x_2=0$, which is of the type $F_2=0$ with $n=2$ and $B=0$, has an integrand function of the form $f=p_1p_2/2+\partial\omega_1/\partial x_1+\partial\omega_2/\partial x_2$ where ω_1, ω_2 are arbitrary functions of x_1, x_2, z.

If $\mathfrak{a}=0$ our discussion proceeds on the basis of the system (2.8) written for the equation $F_2=0$. It is evident that in every case the resulting system of equations S can be replaced by an equivalent system S^* of $R\leqq 2n+1$ independent partial differential equations.[8] The number R of equations in S^* is equal to the rank of (A), the matrix of coefficients of the system S. Inspection of (A) discloses that its last column is a linear combination of the n preceding columns. Thus R is always less than $2n+1$. Actually, when $\mathfrak{a}=0$, R is at most $2n-1$ and its precise value is given by the following lemmas which we state without proof.[9]

[8] Cf. E. Goursat, *Leçons sur l'Intégration des Équations aux Dérivées Partielles du Premier Ordre*, 1921, p. 66.

[9] These proofs and other details omitted in this paper are to be found in the writer's doctoral dissertation, *Inverse Problems of the Calculus of Variations for Multiple Integrals*, The Ohio State University, 1936. For assistance in the preparation of this dissertation and the present paper, the writer wishes to express his indebtedness to Professor Lincoln LaPaz of The Ohio State University.

LEMMA 1. *If* (A_{ij}) *is of rank* r $(r=2, 3, \cdots, n-1)$, *the matrix* (A) *is of rank* $n+r$.

LEMMA 2. *If* (A_{ij}) *is of rank* 1, *the matrix* (A) *is of rank* n.

Where it is desirable to retain the original variables x_i, a straightforward but somewhat involved discussion based on Lemmas 1 and 2 enables us to treat[10] both the case $1<r<n$ and the case $r=1$. We give a simpler treatment of the cases $1\leqq r<n$ using another lemma.

LEMMA 3. *If* (A_{ij}) *is of rank* r, *then by means of a nonsingular linear transformation* $X_i=c_{i\alpha}x_\alpha$ $(i=1, 2, \cdots, n)$, *where the* c_{ij} *are constants, the partial differential equation* $F_2=0$ *can be transformed into one such that*[11] $A_{ij}=\delta_j^i$, $A_{ii}=\pm 1$ $(i\leqq r)$, $A_{ii}=0$ $(i>r)$.

By use of Lemma 3, (2.8) can be replaced by the equivalent system

$$(4.2)\quad A_{ii}\partial M/\partial p_k - A_{ki}\partial M/\partial p_i = 0,\quad i=1, 2, \cdots, r;\ k=1, 2, \cdots, n,$$

$$(4.3)\quad A_{ii}(\partial M/\partial x_i + p_i\partial M/\partial z) - B\partial M/\partial p_i = 0,\qquad i=1, 2, \cdots, n.$$

From (4.2) with $i=1$ and $k>1$ we infer $\partial M/\partial p_k=0$ $(k>1)$. If now $1<r<n$, the choice $i=2$, $k=1$ shows $\partial M/\partial p_1=0$. Hence for $1<r<n$ it follows from $\partial M/\partial p_i=0$ $(i=1, 2, \cdots, n)$ and (4.3) that $\partial M/\partial x_i+p_i\partial M/\partial z=0$ $(i\leqq r)$ and therefore that $\partial M/\partial x_i=\partial M/\partial z=0$ $(i\leqq r)$. This completes the proof of the following theorem.

THEOREM 4.2. *If the rank of* (A_{ij}) *is* r $(1<r<n)$, *then the partial differential equation* $F_2=0$, *when reduced to the normal form of Lemma* 3, *admits as its most general multiplier the function* $A(x_{r+1}, x_{r+2}, \cdots, x_n)$ *where* A *is different from zero but is otherwise an arbitrary function of its arguments.*

We now proceed to calculate the function f of (2.2). A particular solution of (2.3) formed for the equation $A(x_{r+1}, x_{r+2}, \cdots, x_n)\cdot F_2=0$ is $G=\frac{1}{2}A(x_{r+1}, x_{r+2}, \cdots, x_n)A_{\mu\mu}p_\mu p_\mu$. Substituting this value of G in (2.4), we have

$$C_z - D_{\alpha z_\alpha} = -B\cdot A(x_{r+1}, x_{r+2}, \cdots, x_n).$$

A particular solution of this equation is $D_i=0$ $(i=1, 2, \cdots, n)$, $C=-A(x_{r+1}, x_{r+2}, \cdots, x_n)\cdot\int_{z_0}^{z}Bdz$. We thus get the following result.

THEOREM 4.3. *If the rank of* A_{ij} *is* r $(1<r<n)$, *then the most general variation problem* (I) *associated with the equation* $F_2=0$, *when reduced to the normal form of Lemma* 3, *has an integrand function* f *of the form*

[10] See the dissertation cited in footnote 9.

[11] Cf. L. E. Dickson, *Modern Algebraic Theories*, 1926, pp. 71–72.

$$f = A(x_{r+1}, x_{r+2}, \cdots, x_n)\cdot\left[\frac{A_{\mu\alpha}p_\mu p_\alpha}{2} - \int_{z_0}^{z} B dz\right] + \partial\omega_\alpha/\partial x_\alpha$$

where the ω_i are arbitrary functions of $x_1, x_2, \cdots, x_n, z$.

Finally if the rank of (A_{ij}) is $r=1$, then $\partial M/\partial p_k=0$ for all $k>1$ and (4.3) reduces to the single equation $A_{11}(\partial M/\partial x_1+p_1\partial M/\partial z) = B\cdot\partial M/\partial p_1$, which obviously has two independent solutions $M_1(x_1, z, p_1)$ and $M_2(x_1, z, p_1)$ for M. In this case we are thus led to a somewhat different result.

THEOREM 4.4. *If the rank of (A_{ij}) is 1, then the partial differential equation $F_2=0$, when reduced to the normal form of Lemma 3, admits as its most general multiplier the functions $A(M_1, M_2, x_2, x_3, \cdots, x_n)$ where A is different from zero but is otherwise an arbitrary function.*

The determination of the most general integrand function f when $r=1$ may be left to the reader.

As an application of this case we find that the most general multiplier for $r=\partial^2 z/\partial x_1^2=0$, which is of type $F_2=0$ with $n=2$ and $B=0$, is $A(p_1, z-p_1x_1, x_2)$. The most general variation problem (I) associated with $r=0$ has an integrand function f of the form

$$f = \int_{p_{1_0}}^{p_1}\int_{p_{1_0}}^{p_1} A(p_1, z - p_1x_1, x_2)dp_1dp_1 + \partial\omega_1/\partial x_1 + \partial\omega_2/\partial x_2$$

where ω_1, ω_2 are arbitrary functions of x_1, x_2, z.

5. **Variation problems associated with the equation $R(p, q)r+2S(p, q)s+T(p, q)t=0$.** Consider the partial differential equation

$$F_3 \equiv R(p, q)r + 2S(p, q)s + T(p, q)t = 0, \tag{5.1}$$

where $z=z(x, y)$, $p=\partial z/\partial x$, $q=\partial z/\partial y$, $r=\partial^2 z/\partial x^2$, $s=\partial^2 z/\partial x\partial y$, $t=\partial^2 z/\partial y^2$, and the coefficients R, S, T are analytic functions of p and q. We shall exclude equations for which two of the coefficients R, S, T vanish identically, since such vanishing would reduce the equation $F_3=0$ to one of the special equations $r=0$, $s=0$ treated in the previous section. The system (2.7) with $n=2$ and suitable identification of the variables now has the form

$$\begin{aligned}
S\cdot\partial M/\partial p - R\cdot\partial M/\partial q + M\cdot[\partial S/\partial p - \partial R/\partial q] &= 0,\\
T\cdot\partial M/\partial p - S\cdot\partial M/\partial q + M\cdot[\partial T/\partial p - \partial S/\partial q] &= 0,\\
R\cdot[\partial M/\partial x + p\partial M/\partial z] + S\cdot[\partial M/\partial y + q\partial M/\partial z] &= 0,\\
S\cdot[\partial M/\partial x + p\partial M/\partial z] + T\cdot[\partial M/\partial y + q\partial M/\partial z] &= 0.
\end{aligned} \tag{5.2}$$

In contradistinction to (4.1), (5.1) may (Case I) or may not (Case II) have a self-adjoint equation of variation. In Case I, since $\partial S/\partial p - \partial R/\partial q = \partial T/\partial p - \partial S/\partial q = 0$, (5.2) becomes the homogeneous system

$$
\begin{aligned}
&U_1M \equiv S\cdot\partial M/\partial p - R\cdot\partial M/\partial q = 0,\\
&U_2M \equiv T\cdot\partial M/\partial p - S\cdot\partial M/\partial q = 0,\\
&U_3M \equiv R\cdot[\partial M/\partial x + p\cdot\partial M/\partial z] + S\cdot[\partial M/\partial y + q\cdot\partial M/\partial z] = 0,\\
&U_4M \equiv S\cdot[\partial M/\partial x + p\cdot\partial M/\partial z] + T\cdot[\partial M/\partial y + q\cdot\partial M/\partial z] = 0.
\end{aligned}
\tag{5.3}
$$

In either case the equations of (5.2) and (5.3) are found to be independent if and only if $\delta \equiv RT - S^2 \neq 0$. Each of the Cases I, II may therefore be further subdivided according as (1) $\delta \not\equiv 0$, (2) $\delta \equiv 0$. We proceed to a discussion of these cases.

Case I_1. When (5.1) has a self-adjoint equation of variation and $\delta \not\equiv 0$, it follows from Theorem 2.1 that $F_3 = 0$ has a constant as its most general multiplier. Hence we have proved the following.

THEOREM 5.1. *If the equation $F_3 = 0$ has a self-adjoint equation of variation in a neighborhood in which $\delta \neq 0$, then the most general variation problem* (I) *associated with $F_3 = 0$ in this neighborhood has an integrand function f of the form $f = G + \partial\omega_1/\partial x + \partial\omega_2/\partial y$, where $G(x, y, z, p, q)$ is a particular solution of $g_{pp} = R$, $g_{pq} = S$, $g_{qq} = T$, satisfying $G_{pz} + G_{qy} + pG_{pz} + qG_{qz} - G_z = 0$ and the ω_i are arbitrary functions of x, y, z.*

As an illustration we note that the most general variation problem (I) associated with the equation $q^2r + 4pqs + p^2t = 0$ has an integrand function of the form $f = p^2q^2/2 + \partial\omega_1/\partial x + \partial\omega_2/\partial y$, where the ω_i are arbitrary functions of x, y, z.

Case I_2. When $\delta \equiv 0$ it follows that $R\cdot S\cdot T \not\equiv 0$ and hence, in any neighborhood in which $R\cdot S\cdot T \neq 0$, $U_{k+1}M \equiv (T/S)\cdot U_kM$ $(k = 1, 3)$. The system $U_iM = 0$ $(i = 1, 3)$ is therefore equivalent to (5.3) and moreover is found to be complete. Hence we have the following result.

THEOREM 5.2. *If $F_3 = 0$ has a self-adjoint equation of variation in a neighborhood in which $R\cdot S\cdot T \neq 0$ and $\delta \equiv 0$, then it admits as its most general multiplier in this neighborhood the function $M(x, y, z, p, q) = A(M_1, M_2, M_3)$ where M_i $(i = 1, 2, 3)$ are any three independent integrals of $U_iM = 0$ $(i = 1, 3)$ and $A \neq 0$ is an arbitrary function.*

The most general variation problem (I) can now be obtained in the usual manner. The equation $\phi(p+q)\cdot[r + 2s + t] = 0$, where ϕ is an arbitrary analytic function of its argument, illustrates this case.

We now take up Case II in which $F_3 = 0$, as it stands, does not have a self-adjoint equation of variation. If in addition $\delta \neq 0$, then we can

infer from Theorem 2.2 that at most one multiplier $M(x, y, z, p, q)$ exists. To secure more precise information in this case we must discuss the nonhomogeneous system (5.2). To integrate (5.2) we transform it into a system linear and homogeneous in the first partial derivatives of a function $m(x, y, z, p, q, M)$ with $\partial m/\partial M \neq 0$, which defines M by means of the relation $m = \text{const}$. The resulting system

$$
\begin{aligned}
U_1 m &\equiv S\partial m/\partial p - R\partial m/\partial q - M\cdot[\partial S/\partial p - \partial R/\partial q]\cdot\partial m/\partial M = 0,\\
U_2 m &\equiv T\partial m/\partial p - S\partial m/\partial q - M\cdot[\partial T/\partial p - \partial S/\partial q]\cdot\partial m/\partial M = 0,\\
U_3 m &\equiv R(\partial m/\partial x + p\partial m/\partial z) + S(\partial m/\partial y + q\partial m/\partial z) = 0,\\
U_4 m &\equiv S(\partial m/\partial x + p\partial m/\partial z) + T(\partial m/\partial y + q\partial m/\partial z) = 0
\end{aligned}
\tag{5.4}
$$

is independent since $\delta \neq 0$. It is easy to verify that the commutator $(U_1U_4)m \equiv U_5 m$ is not a linear combination of $U_i m$ $(i = 1, 2, 3, 4)$ and we accordingly adjoin $U_5 m = 0$ to the system (5.4). It is next found that the commutator $(U_1U_2)m$ is a linear combination of $U_i m$ $(i = 1, 2, 3, 4, 5)$ if and only if the following determinant D vanishes:

$$
\begin{vmatrix}
S & T & S\cdot[\partial T/\partial p + \partial S/\partial q] - R\cdot\partial T/\partial q - T\cdot\partial S/\partial p\\
R & S & S\cdot[\partial S/\partial p + \partial R/\partial q] - R\cdot\partial S/\partial q - T\cdot\partial R/\partial p\\
\partial S/\partial p - \partial R/\partial q & \partial T/\partial p - \partial S/\partial q & V
\end{vmatrix}
$$

where

$$
\begin{aligned}
V = {} & S\cdot[\partial^2 T/\partial p^2 - \partial^2 R/\partial q^2] - R\cdot[\partial^2 T/\partial p\partial q - \partial^2 S/\partial q^2]\\
& - T\cdot[\partial^2 S/\partial p^2 - \partial^2 R/\partial q\partial p].
\end{aligned}
$$

We are thus led to distinguish two subcases II_{11} and II_{12} according as $D \neq 0$ or $D \equiv 0$.

Case II_{11}. In a neighborhood in which $\delta \neq 0$ and $D \neq 0$, $(U_1U_2)m \equiv U_6 m$ is not a linear combination of $U_i m$ $(i = 1, 2, 3, 4, 5)$, and upon adjoining $U_6 m = 0$ to our system we have a system of six independent equations in the partial derivatives of $m(x, y, z, p, q, M)$. Hence the most general integral of the system is a constant. Thus there exists no nonsingular solution of the system (5.2) and we have the following.

THEOREM 5.3. *If $F_3 = 0$ does not have a self-adjoint equation of variation in a neighborhood in which $\delta \neq 0$ and $D \neq 0$, it admits no multiplier in this neighborhood.*

The equation $p^2 r + 2q^2 s = 0$ is an illustration of this case.

Case II_{12}. In a neighborhood in which $\delta \neq 0$ and $D \equiv 0$ the system $U_i m = 0$ $(i = 1, 2, 3, 4, 5)$ is complete and has $m = A(m_1)$ for its most general integral where $m_1(p, q, M)$ is a particular solution of $U_i m = 0$

($i=1, 2$) and A is an arbitrary function of its argument. The most general nonsingular solution M of the system (5.2) is obtained by solving for M the relation $A(m_1)=\text{const.}$ The equation $M\cdot F_3=0$ now has a self-adjoint equation of variation and $M^2\cdot\delta\neq 0$. Hence, since multipliers which differ by a nonzero constant factor are not regarded as distinct, we have from Theorem 2.1 the following.

THEOREM 5.4. *If $F_3=0$ does not have a self-adjoint equation of variation in a neighborhood in which $\delta\neq 0$ and $D\equiv 0$, then it admits as its most general multiplier in this neighborhood the unique solution for M of the relation $A(m_1)=\text{const.}$*

The most general variation problem (I) can now be obtained in the usual manner. As an illustration of this case we cite the partial differential equation of minimal surfaces[12] $(1+q^2)r-2pqs+(1+p^2)t=0$.

Case II_2. If $\delta\equiv 0$ it follows that $R\cdot S\cdot T\not\equiv 0$, and hence in any neighborhood in which $R\cdot S\cdot T\neq 0$ we find that $U_4m\equiv(T/S)\cdot U_3m$. However, $U_2m\equiv(T/S)\cdot U_1m$ if and only if $\Delta\equiv\partial(R/S)^2/\partial q-2\partial(R/S)/\partial p$ vanishes. We therefore consider the two subcases II_{21} and II_{22} according as $\Delta\equiv 0$ or $\Delta\not\equiv 0$.

Case II_{21}. In a neighborhood in which $R\cdot S\cdot T\neq 0$ and $\Delta\equiv 0$, the system $U_im=0$ ($i=1, 3$) is equivalent to (5.4) and furthermore is complete. Hence we have the following result in this case.

THEOREM 5.5. *If $F_3=0$ does not have a self-adjoint equation of variation in a neighborhood in which $R\cdot S\cdot T\neq 0$, $\delta\equiv 0$ and $\Delta\equiv 0$, then in this neighborhood $F_3=0$ admits as its most general multiplier the solution for M of the relation $A(m_1, m_2, m_3, m_4)=\text{const.}$, where m_i ($i=1, 2, 3, 4$) are any four particular independent solutions of $U_im=0$ ($i=1, 3$) and where A is different from zero but is otherwise an arbitrary function.*

The equation $q^2r-2pqs+p^2t=0$ is an illustration of this case.

Case II_{22}. In a neighborhood in which $R\cdot S\cdot T\neq 0$ and $\Delta\neq 0$ we find it more perspicuous to consider the system (5.2). If in (5.2) we multiply the first equation by $S(p, q)$, the second by $R(p, q)$ and subtract, we obtain $M\cdot[S\cdot\partial S/\partial p-S\cdot\partial R/\partial q+R\cdot\partial S/\partial q-R\cdot\partial T/\partial p]=0$. Since $\Delta\neq 0$ we must have $M=0$, and hence the following is valid.

THEOREM 5.6. *If $F_3=0$ does not have a self-adjoint equation of variation in a neighborhood in which $R\cdot S\cdot T\neq 0$, $\delta\equiv 0$ and $\Delta\neq 0$, then it admits no multiplier in this neighborhood.*

The equation $p^2r+2pqs+q^2t=0$ illustrates this case.

FENN COLLEGE

[12] Cf. L. LaPaz, Acta Szeged, vol. 5 (1932), pp. 199–207.

COPELAND'S DEFINITION OF A STIELTJES INTEGRAL

R. L. JEFFERY

Our present interest in Copeland's definition[1] of the Stieltjes integral of $g(x)$ with respect to the monotone function $f(x)$ is due to a remark by T. H. Hildebrandt[2] to the effect that in the proof of the formula for integration by parts it is required that $f(x)=\frac{1}{2}\{f(x+0)+f(x-0)\}$. In looking over Copeland's paper it was found that it is only in the proof of this formula that $f(x)$ is so restricted. Furthermore, it became clear that the definition possesses a considerable degree of generality. In the present note we simplify the definition, and compare it with that of the Riemann-Stieltjes integral and the Lebesgue-Stieltjes integral.

The classical definition of a Riemann-Stieltjes integral is

$$\text{(1)} \qquad \text{RS} \int_\alpha^\beta g\,df = \lim_{n\to\infty} \sum_{k=1}^{n} g(\xi_k)\{f(x_k) - f(x_{k-1})\},$$

where (x_{k-1}, x_k) is a finite subdivision of the interval $\alpha \leqq x \leqq \beta$, with $x_i - x_{i-1} \to 0$ and ξ_k any point on (x_{k-1}, x_k). The limit (1) exists when g is continuous, but may fail to exist even for functions g of bounded variation unless further restrictions[3] are placed on the subdivision (x_{k-1}, x_k) or on the choice of ξ_k. Copeland's definition is likewise based on what can be interpreted as a sequence of finite sets $\{(x_k)\}$, $k=1, 2, \cdots, n$, of (α, β), and the integral is given by

$$\text{(2)} \qquad \text{CS} \int_\alpha^\beta g\,df = \lim_{n\to\infty} \frac{g(x_1) + \cdots + g(x_n)}{n}.$$

The sequence $\{x_k\}$ is defined wholly in terms of f, and the limit (2) exists for a wide class of functions including functions of bounded variation.

The set $\{x_k\}$, $k=1, 2, \cdots, n$; $n=1, 2, \cdots$, on which (2) is based is denumerable. Consequently the value of the integral depends only on the values of g over this denumerable set, which permits g an undesirable amount of freedom. To obviate this defect, and to bring the definition more in line with that of the Riemann-Stieltjes integral, we introduce some changes in its formulation.

On the interval $\alpha \leqq x \leqq \beta$ let $g(x)$ be bounded, and $f(x)$ be bounded

[1] This Bulletin, vol. 43 (1937), pp. 581–588.

[2] American Mathematical Monthly, vol. 45 (1938), p. 277.

[3] This point has been thoroughly covered by Hildebrandt, loc. cit., §§6, 7, 8, 9.

and non-decreasing. In order to keep the initial stages as simple as possible, we shall assume that $f(\alpha+0)=0$, $f(\beta-0)=1$. This restriction will be removed later. For a given n and for $k=1, 2, \cdots, n-1$ let x_k be the greatest lower bound of numbers x for which

$$f(x-0) \leqq k/n \leqq f(x+0). \tag{3}$$

The numbers x_k, $(k=1, 2, \cdots, n-1)$, all lie on the open interval $\alpha<x<\beta$. Let $a<x<b$ be an open interval on $\alpha<x<\beta$, and for any n let p be the number of points of the set $\{x_k\}$ on $a<x<b$. Then

$$f(b-0)-f(a+0)-1/n<p/n<f(b-0)-f(a+0)+1/n,$$

from which it follows that

$$\lim_{n\to\infty} p/n = f(b-0)-f(a+0). \tag{4}$$

Also if x_0 is a point of discontinuity of f on $\alpha<x<\beta$ and q is the number of times x_0 is repeated in the set $\{x_k\}$, then

$$f(x_0-0)-1/n<q/n<f(x_0+0)+1/n,$$

which gives

$$\lim_{n\to\infty} q/n = f(x_0+0)-f(x_0-0). \tag{5}$$

The set $\{x_k\}$, $k=1, 2, \cdots, n-1$, is non-decreasing on $\alpha<x<\beta$. Let ξ_1 be a point on $\alpha<x<x_1$, ξ_k a point on $x_{k-1}\leqq x\leqq x_k$ if $x_k\neq x_{k-1}$, $\xi_k=x_k$ if $x_k=x_{k-1}$, $k=2, 3, \cdots, n-1$, and ξ_n a point on $x_{n-1}<x<\beta$. Form the sum

$$G_n=\frac{g(\xi_1)+\cdots+g(\xi_n)}{n},$$

and over the open interval $\alpha<x<\beta$ define

$$\mathrm{CS}\int_{\alpha<x<\beta} g\,df = \lim_{n\to\infty} G_n \tag{6}$$

provided this limit exists. If x_0 is a point of discontinuity of f on $\alpha<x<\beta$, it follows from (5) that the part of G_n arising from x_0 tends to $g(x_0)\{f(x_0+0)-f(x_0-0)\}$. This leads us to define

$$\mathrm{CS}\int_{x_0} g\,df = g(x_0)\{f(x_0+0)-f(x_0-0)\}, \tag{7}$$

which may be retained when x_0 is a point of continuity of f. If $a<x<b$

is an open interval on $\alpha<x<\beta$, set $g'=g$ on $a<x<b$, $g'=0$ elsewhere on $\alpha<x<\beta$, and define

$$\text{(8)} \qquad \text{CS}\int_{\alpha<x<b} g\,df = \text{CS}\int_{\alpha<x<\beta} g'\,df.$$

We further define

$$\text{CS}\int_a^b g\,df = \text{CS}\int_{a<x<b} g\,df + g(a)\{f(a+0)-f(a)\} + g(b)\{f(b)-f(b-0)\}.$$

From (7) and (8) we see that

$$\text{CS}\int_a^b g\,df = \int_{a<x<b} g\,df + \int_a g\,df + \int_b g\,df$$

only if $f(a)=f(a-0)$, and $f(b)=f(b+0)$. It is true, however, that if c is a point on $a<x<b$, then

$$\text{CS}\int_a^b g\,df = \text{CS}\int_a^c g\,df + \text{CS}\int_c^b g\,df.$$

With the restrictions $f(\alpha+0)=0$, $f(\beta-0)=1$ still holding, the Copeland-Stieltjes integral, $\text{CS}\int_a^b g\,df$, has now been defined for every interval (a, b) on $\alpha\leqq x\leqq\beta$. We shall prove the following theorems.

THEOREM I. *The necessary and sufficient condition for the existence of the* CS*-integral is that the common part of the discontinuities of g and the continuities of f have zero measure with respect to f.*

THEOREM II. *If the* CS*-integral exists, the* LS*-integral (Lebesgue-Stieltjes integral) exists, and the two are equal.*

THEOREM III. *If* $F(x)=\int_a^x g\,df$, *then* dF/df *exists and is equal to g except for at most a set of zero measure with respect to f.*

It is thus seen that CS-integration with respect to the monotone function $f(x)$ has the same degree of generality as ordinary Riemann integration with respect to the variable x. In other words: If Copeland's definition replaces that of Riemann, it makes no difference whether the variable of integration is x or a monotone function $f(x)$.

Let $D=d_1, d_2, \cdots$ be the points of discontinuity of f, and for any point d_i let (a_i, a_i') be an interval with $a_i<d_i<a_i'$ and for which a_i and a_i' are points of continuity of f. Let p_i, q_i, r_i be respectively the number of points of the set $\{x_k\}$ on $a_i<x_i<d_i$, the number of times d_i

is repeated in this set, and the number of points of the set on $d_i<x<a_i'$. Then $p_i/n \to f(d_i-0)-f(a_i)$, $q_i/n \to f(d_i+0)-f(d_i-0)$, and $r_i/n \to f(a_i')-f(d_i+0)$. Thus if $a_i'-a_i$ is sufficiently small and n is sufficiently great, the part of G_n arising from the interval (a_i, a_i') is arbitrarily near $g(d_i)\{f(d_i+0)-f(d_i-0)\} = \text{CS}\int_{d_i} g df$. Since g is bounded and f is bounded and non-decreasing, $\sum g(d_i)\{f(d_i+0)-f(d_i-0)\}$ converges. Also, there exists a set of nonoverlapping intervals $A=(a_1, a_1'), \cdots, (a_l, a_l')$ with $a_i<d_i<a_i'$ and such that for l sufficiently great and n sufficiently great we have

$$\sum_{i=l+1}^{\infty} f(d_i+0)-f(d_i-0)<\epsilon, \tag{9}$$

and, on account of relations (4) and (5), the part of G_n arising from the intervals A differs from

$$\sum_{i=1}^{l} g(d_i)\{f(d_i+0)-f(d_i-0)\} = \sum \int_{d_i} g df$$

by not more than ϵ. Hence the upper and lower limits of the part of G_n arising from the intervals A do not differ by more than 2ϵ.

Let e_ϵ be the set of points of (α, β) at which the saltus of g is not less than ϵ. Then e_ϵ is closed, and consequently the part e_ϵ' of e_ϵ which is not interior to A is closed. Let the common part of e_ϵ and the set of continuities of f have zero measure with respect to f. It then follows, if (9) is taken into consideration, that the closed set e_ϵ' can be put in a finite set of nonoverlapping intervals $B=(b_i, b_i')$ which do not overlap the set A, which are such that b_i, b_i' are points of continuity of f, and for which

$$\sum \{f(b_i')-f(b_i)\}<2\epsilon.$$

It then follows that for any n the part of G_n arising from the intervals B is not greater in numerical value than $2\epsilon M$, where M is the least upper bound of $|g(x)|$.

At each point of (α, β) not interior to the intervals $A+B$ the saltus of g is less than ϵ. Hence the closed intervals complementary to the set interior to $A+B$ can be subdivided into a finite set of intervals $C=(c_i, c_i')$ such that c_i, c_i' are points of continuity of f and on each interval (c_i, c_i') the fluctuation of g is not more than ϵ. If $G_n'=\{g(\xi_1')+\cdots+g(\xi_m')\}/n$ is the part of G_n arising from the interval (c_i, c_i'), then G_n' lies between the values

$$\{g(c_i)-\epsilon\}p_i/n, \qquad \{g(c_i)+\epsilon\}p_i/n,$$

where p_i is the number of points of the sequence $\{x_k\}$ on $c_i<x<c_i'$. Hence as $n\to\infty$ the upper and lower limits of G_n' do not differ by more than $2\epsilon\{f(c_i')-f(c_i)\}$. We can now conclude that the least upper bound and greatest lower bound of

$$G_n=\frac{g(\xi_1)+\cdots+g(\xi_n)}{n}$$

do not differ by more than $2\epsilon+2\epsilon M+2\epsilon\{f(\beta)-f(\alpha)\}$. Since ϵ is arbitrary, it follows that if $g(x)$ satisfies the conditions of Theorem I then the CS-integral of g with respect to the monotone function $f(x)$ exists.

If the set of discontinuities of g has measure greater than 0 with respect to f over the continuities of f, then for some $\epsilon>0$ there exists $d>0$ such that $\sum\{f(b_i')-f(b_i)\}>d$ regardless of the choice of the intervals (b_i, b_i'). Consequently for every n there exist ξ_i and ξ_i' such that the sums $\{g(\xi_1)+\cdots+g(\xi_n)\}/n$ and $\{g(\xi_1')+\cdots+g(\xi_n')\}/n$ will differ by an amount which, for the fixed ϵ in question, is not less than ϵd. From this it follows that the CS-integral of f with respect to g does not exist, and we conclude that the conditions of Theorem I are necessary.[4]

To obtain the LS-integral of g with respect to f we proceed as follows: Let (y_{i-1}, y_i) be a subdivision of the range of g. Let e_i be the part of (α, β) for which $y_{i-1}\leqq g<y_i$, and $f(e_i)$ the image of the set e_i by means of the transformation $y=f(x)$, where a point x_0 of discontinuity of f corresponds to the closed interval $\{f(x_0-0), f(x_0+0)\}$. Then $\mathrm{LS}\int_\alpha^\beta g df=\lim\sum y_i\, mf(e_i)$ as $y_i-y_{i-1}\to 0$, provided this limit exists. This limit does exist if g is measurable with respect to f, which readily follows if the points of continuity of f at which g is discontinuous have zero measure with respect to f. If $D=d_1, d_2, \cdots$ is the set of discontinuities of f, then

$$\mathrm{LS}\int_D g df=\sum g(d_i)\{f(d_i+0)-f(d_i-0)\}=\mathrm{CS}\int_D g df.$$

If the intervals $A=(a_i, a_i')$ and $B=(b_i, b_i')$ are properly chosen, then $\mathrm{CS}\int_A g df$ is arbitrarily near to $\sum g(d_i)\{f(d_i+0)-f(d_i-0)\}=\mathrm{LS}\int_D g df$, and both $\mathrm{CS}\int_B g df$ and $\mathrm{LS}\int_B g df$ are less in numerical value than $2\epsilon M$.

[4] The necessity of the conditions of Theorem I would not follow if G_n were restricted to the form $\{g(x_1)+\cdots+g(x_n)\}/n$ of Copeland's original definition. Let $f(x)=x$. The sequence $\{x_k\}$, $k=1, 2, \cdots, n$, $n=1, 2, \cdots$, is a denumerable set on (α, β). If $g(x)=0$ at each point of this denumerable set, and for the rest of (α, β) $g(x)=1$ when x is rational and $g(x)=0$ when x is irrational, then $\mathrm{CS}\int_\alpha^\beta g df$ exists and is equal to zero, but g does not satisfy the conditions of Theorem I.

Over the set of intervals (c_i, c_i') both the CS-integral and the LS-integral of g with respect to f lie between $\sum\{g(c_i)-\epsilon\}\{f(c_i')-f(c_i)\}$ and $\sum\{g(c_i)+\epsilon\}\{f(c_i')-f(c_i)\}$. From these considerations the truth of Theorem II readily follows.

In proceeding with the proof of Theorem III we first define the derivative of a function $F(x)$ with respect to a monotone function $f(x)$. Let

$$\psi(x, \Delta x) = \frac{F(x+\Delta x) - F(x-0)}{f(x+\Delta x) - f(x-0)}, \qquad f(x+\Delta x) - f(x-0) \neq 0, \Delta x > 0,$$

$$= \frac{F(x+\Delta x) - F(x+0)}{f(x+\Delta x) - f(x+0)}, \qquad f(x+\Delta x) - f(x+0) \neq 0, \Delta x < 0,$$

$$= 0 \text{ otherwise.}$$

If, for a fixed x, $\psi(x, \Delta x)$ tends to a limit as $\Delta x \to 0$, then this limit is the derivative of $F(x)$ with respect to $f(x)$, dF/df. Now let $F(x) = \int_a^x g df$. If x_0 is a point of discontinuity of f, then, for $\Delta x > 0$, $F(x_0+\Delta x) - F(x_0-0)$ is the limit as $\delta \to 0$ of $F(x_0+\Delta x) - F(x_0-\delta)$, and this last is the limit of

$$\frac{g(\xi_1^-) + \cdots + g(\xi_p^-) + qg(x_0) + g(\xi_1^+) + \cdots + g(\xi_r^+)}{n} \tag{10}$$

where $\xi_1^-, \cdots, \xi_p^-, \xi_1^+, \cdots, \xi_r^+$ are points on the intervals (x_{k-1}, x_k) formed by the sequence $\{x_k\}$ on $x_0-\delta < x < x_0$, $x_0 < x < x_0+\Delta x$ respectively, and q is the number of times x_0 is repeated in this sequence. Then for δ and Δx fixed $p/n \to f(x_0-0) - f(x_0-\delta+0)$, $q/n \to f(x_0+0) - f(x_0-0)$, and $r/n \to f(x_0+\Delta x-0) - f(x_0+0)$. Hence for δ and Δx sufficiently small p/n and r/n are both arbitrarily near to zero. Since g is bounded, it then follows from (10) that for δ and Δx sufficiently small $F(x_0+\Delta x) - F(x_0-\delta)$ is arbitrarily near to $g(x_0)\{f(x_0+0) - f(x_0-0)\}$, and consequently $\psi(x_0, \Delta x)$ is arbitrarily near to $g(x_0)$. Similarly it can be shown that for $\Delta x < 0$, $\psi(x_0, \Delta x) \to g(x_0)$.

Next let x_0 be a point of continuity of both g and f. Then $F(x_0+\Delta x) - F(x_0-\delta)$ is the limit of a sum of the form

$$\frac{g(x_0) + t_1 + \cdots + g(x_0) + t_p}{n}$$

where $|t_i| < \epsilon$ if δ and Δx are sufficiently small, and $p/n \to f(x_0+\Delta x-0)$

$-f(x_0-\delta+0)$. Hence for δ and Δx sufficiently small, and n sufficiently great

$$\psi(x_0, \Delta x) = g(x_0) + t\{f(x_0 + \Delta x - 0) - f(x_0 - \delta + 0)\}$$

where $t\to 0$ as Δx and $\delta\to 0$. We conclude therefore that, for $\Delta x>0$, $\psi(x_0, \Delta x)\to g(x_0)$. Similarly it can be shown that, for $\Delta x<0$, $\psi(x, \Delta x)\to g(x_0)$. Hence, at x_0, $dF/df=g(x_0)$, and Theorem III is established.

We note that the definition of dF/df given above implies the existence of $F(x-0)$ and $F(x+0)$, which can easily be shown if $F(x)=\mathrm{CS}\int_a^x g df$. Furthermore, at points of discontinuity of $F(x)$ or of $f(x)$ the value of the derivative does not depend on the value of these functions at the point.

The next considerations are the removal of the restrictions $f(\alpha+0)=0$, $f(\beta-0)=1$, and the equivalence of the original and the modified definition. As to the first of these, it can be done, following Copeland, by setting

$$\phi(x) = \frac{f(x) - f(\alpha + 0)}{f(\beta - 0) - f(\alpha + 0)}$$

and defining

$$(11) \qquad \mathrm{CS} \int_{\alpha<x<\beta} g df = \mathrm{CS} \int_{\alpha<x<\beta} g d\phi \cdot [f(\beta - 0) - f(\alpha + 0)].$$

It is possible, however, and perhaps advisable, to forego this restriction from the start: In (3) let x_k be the greatest lower bound of numbers x for which

$$f(x - 0) \leqq f(\alpha + 0) + (k/n)\{f(\beta - 0) - f(\alpha + 0)\} \leqq f(x + 0),$$

and define

$$(12) \qquad \mathrm{CS} \int_{\alpha<x<\beta} g df = \lim_{n\to\infty} \frac{g(\xi_1) + g(\xi_2) + \cdots + g(\xi_n)}{\sigma_n},$$

where ξ_k is chosen as above, and $1/\sigma_n=\{f(\beta-0)-f(\alpha+0)\}/n$. If $1/\sigma_n$ replaces $1/n$ in the foregoing discussions, all the results hold without further change in the wording. When $f(\alpha+0)=0$ and $f(\beta-0)=1$, (12) reduces to (6); and when these restrictions do not hold, it can be shown that (11) and (12) are equivalent. The proof of this we leave to the reader.

As to the equivalence of the modified form and the original form, if the integral exists under the first it does under the second. For the

part contributed by the discontinuities of f is the same; on the intervals (c_i, c_i') both integrals lie between the values

$$(13) \quad \sum g(c_i)\{f(c_i') - f(c_i)\} - \epsilon K, \quad \sum g(c_i)\{f(c_i') - f(c_i)\} + \epsilon K,$$

where $K = f(\beta) - f(\alpha)$. These considerations easily lead to the equality of the two. That the integral can exist under the original form and not under the modified form is shown by the example in the footnote above. The set[5] $x_1; x_2, x_3; x_4, x_5, x_6; \cdots$ on which the original definition is based is such that for any n the first n points are not arranged in increasing order of magnitude. But obviously these can be rearranged into a non-decreasing set $\{x_k\}$, and on (x_{k-1}, x_k) the ξ_k can be chosen as above. Then

$$\lim_{n\to\infty} \frac{g(\xi_1) + \cdots + g(\xi_n)}{n}$$

leads to an integral which is equivalent to that arising from the modified definition given in this paper. For the results and proof of Theorem I are the same; the part of each arising from the discontinuities of f is the same; furthermore, on the intervals (c_i, c_i') both integrals lie between the bounds given by (13).

Thus, if we get away from an integral which depends only on the values of g over a denumerable set, the sequence of sets given by (3) leads to the same result as the set $x_1; x_2, x_3; \cdots$ on which the original definition is based. The sequence given by (3) is more simply defined; the two are equivalent with respect to properties (4) and (5), but for that given by (3) these properties are more easily established.

In correspondence Copeland has stated that he had in mind statistical considerations when he developed the original definition. "These numbers," that is, the set $x_1; x_2, x_3; \cdots$, "can be interpreted as a sequence of measurements of some physical quantity. The expression $f(b+0) - f(a+0)$ is interpreted as the probability that a measurement x_k will lie in the interval $a < x \leqq b$. This probability is defined as the limit of the success ratio, that is, as the average number of points lying in the interval. If a person is to receive $g(x)$ dollars when the measurement turns out to be x, then the expected amount which he will receive is the limit of the average of his receipts. This limit is the integral $\int_a^\beta g df$." Whatever the utilitarian background of the idea, a Stieltjes integral of the Riemann type applicable without modification to such a wide range of functions is interesting for its own sake.

THE UNIVERSITY OF SASKATCHEWAN

[5] This Bulletin, loc. cit., p. 582.

A THEOREM ON THE ROTATION GROUP OF THE TWO-SPHERE[1]

DEANE MONTGOMERY AND LEO ZIPPIN

Let R be the group of all rotations of euclidean three-dimensional space E about its origin. We shall take the domain of operation of R to be a euclidean two-sphere S with center at the origin. On the space S, R is transitive. The group R is topological and, considered as a space, is homeomorphic to projective three-space.

While studying the action of groups in certain spaces, the following theorem, which, as far as we know, is not in the literature, occurred to us.

THEOREM. *Let G be any proper subgroup of R. Then G is not transitive on S.*

The subgroup G is subject to no restrictions whatever; in particular it is not closed. The proof will be essentially topological and we begin by assuming that G is transitive on S.

The identity element e of G must be a limit point of G. For otherwise G would be finite and hence certainly not transitive on S. Let g be an element of G which is near e. The element g has a pair of fixed points on S one of which will be denoted by p. The element g is a rotation through a small angle A around the line through p and the origin. Under this rotation points near p move in a certain direction around p, say the clockwise direction.

Now let x be any point of S, and let h be any element of G such that $h(p)=x$. The element hgh^{-1} is in G, and it is a rotation through the angle A around the line through x and the origin. Points near x will be moved in a clockwise direction around x.

We have therefore shown that for every point x of the sphere, G contains a rotation through the angle A around the line through the origin and x. For each x, points near x are moved by the associated rotation in a specified sense around x. Let M be the totality of all these rotations the existence of which has just been demonstrated. The set M is homeomorphic to S and consequently is a two-sphere; furthermore M is a subset of G.

Consider the elements of G given by $g^{-1}M$. This set is a two-sphere passing through the identity element. There will be an arc L, which is in $g^{-1}M$ and therefore in G, leading from e to some element of G,

[1] Presented to the Society, October 28, 1939, under the title *Note on rotation-group of the two-sphere.*

call it g' distinct from e. Let U be an open three-cell subset of R which includes e. Let N be a two-sphere made up of elements of G. The existence of arbitrarily small two-spheres of this kind is proved as above by choosing the element g sufficiently near to e, and we may assume that N is in U. We may also assume that LN is in U and that $g'N$ is outside N: as must be the case if N is small enough.

The arc L may now be used to define a deformation of N to $g'N$. Under this deformation all points swept out by N are in G. Furthermore every point inside N is swept out by the deformation. Hence every point of R inside N is in the group G. The group G is thus seen to contain open subsets and, because of homogeneity, G is open in R. It must therefore coincide with R. The assumption that a proper subgroup G was transitive on S has now led to a contradiction, and the proof is therefore complete.

SMITH COLLEGE AND
QUEENS COLLEGE

ON ORDERED ALGEBRAS[1]

A. A. ALBERT

In his first Madison Colloquium lecture M. H. Stone connected the theory of convex bodies with linear sets over an ordered field. It was natural then to ask whether his theory could be extended by replacing these fields by ordered rings and indeed to ask whether there exist ordered rings which are not fields. I discussed this question at that time with S. MacLane and we attempted to answer it. MacLane has since found an example,[2] in the literature, of a noncommutative ordered quasi-field. It is not an algebra (of finite order) however and it is my purpose in this note to give a very brief proof in elementary language of the following decisive result.

THEOREM. *Every ordered algebra is a field.*

We first observe some known consequences of the order postulates.[3] The postulates on products imply that an ordered ring contains no divisors of zero and hence that every ordered algebra is a division algebra D. Then D has a unity quantity $1=1^2>0$, the sums

[1] Presented to the Society, December 2, 1939.

[2] Cf. Reidemeister, *Grundlagen der Geometrie*, p. 40. It is also shown in this text that archimedian ordered quasi-fields are fields.

[3] The order postulates on page 40 of my *Modern Higher Algebra* were called postulates for an ordered field but are valid for arbitrary rings.

$1+1+\cdots+1$ are all positive, D is a normal division algebra of degree n over a non-modular field K, its centrum. If $u \neq 0$ is in D, then $uu^{-1}=1>0$, u and u^{-1} are either both positive or both negative, the transforms udu^{-1} of every positive quantity of D are all positive.

Let then $n>1$ so that D contains a quantity b not in its centrum K. The minimum function of b is $g(x)=x^n+a_1x^{n-1}+\cdots+a_n$ with $n>1$ and the a_i in K. Since K is non-modular, D contains $b-m^{-1}a_1$ and its negative. One of these is a positive quantity d of D with minimum function $h(y)=y^m+b_2y^{m-2}+\cdots+b_m$ for b_i in K. By a theorem of Wedderburn[4] $h(y)=(y-d_m)\cdots(y-d_1)$ for transforms d_i of d, the $d_i>0$, their sum is $d_1+\cdots+d_m=0$ by our choice above. This contradicts our hypothesis that D is ordered. Hence $n=1$, $D=K$ is a field.

In more technical terms our proof is simply the remark that, if it exists, a noncommutative ordered division algebra contains a positive quantity d of zero trace, whereas this trace is a sum of transforms of d and must be positive.

UNIVERSITY OF CHICAGO

[4] See the proof of Theorem 3, p. 230, of L. E. Dickson's *Algebras and their Arithmetics*.

SIMPLICIAL INTERSECTION CHAINS FOR AN ABSTRACT COMPLEX

W. W. FLEXNER

Abstract complexes have been defined by J. W. Alexander, W. Mayer, A. W. Tucker and S. Lefschetz.[1] The definition here adopted is that of Lefschetz which derives directly from that of Tucker. To conform to Lefschetz's present usage the notation here will differ from that in the article cited by having $K=\{E_i^p\}$ and $K^*=\{E_p^i\}$. It should be remarked that when K is infinite $FF=0$ implies that for given E_i^p and E_j^{p-2},

$$[E_i^p : E_k^{p-1}][E_k^{p-1} : E_j^{p-2}] \neq 0$$

for at most a finite number of E_k^{p-1}. The dual K^* here is to have the property

$$[E_i^p : E_j^{p-1}] = (-1)^{p+1}[E_{p-1}^j : E_p^i] = \mu_{ij}^p$$

where μ_{ij}^p is an abbreviated notation (compare loc. cit., p. 346 (c)). This is in order that a fourth postulate

IV. Ki $(E_i^p \cdot E_p^i) = 1$

may be added to I–III (loc. cit., p. 350), where Ki means Kronecker index, the sum of the coefficients of the elements of the zero-chain in question. Here C is a p-chain of K, D a q-chain of K^*, $p>q$.

It is well known that when K is a simplicial or polyhedral manifold, for every pair E_i^p, E_q^j, $p \geqq q$, an intersection can be defined which is a chain of the simplicial regular subdivision of K. And when K is an arbitrary simplicial complex, the intersection is on K and hence simplicial. Here it is shown that *for any complex the intersections $E_i^p \cdot E_q^j$ can be regarded as integral chains of a simplicial complex K^*.*

This complex will be defined abstractly by means of its vertices σ_i^p. The first definition is for $p=q$:

$$E_i^p \cdot E_p^j = \delta_i^j \sigma_j^p \tag{1}$$

(where $\cdot$ replaces the $\odot$ of the article cited). Then for $p>q$, if $r=p-q-1$,

$$E_i^p \cdot E_q^j = \sum_{k_1,k_2,\cdots,k_r} \mu_{ik_1}^{p}\mu_{k_1k_2}^{p-1} \cdots \mu_{k_rj}^{q+1}\sigma_i^p\sigma_{k_1}^{p-1} \cdots \sigma_{k_r}^{q+1}\sigma_j^q \tag{2}$$

[1] S. Lefschetz, this Bulletin, vol. 43 (1937), pp. 345–359. (References to the other authors will be found on page 345.)

where $\sigma_i^p\sigma_{k_1}^{p-1}\cdots\sigma_{k_r}^{q+1}\sigma_j^q$ denotes the simplex of K^r with the given vertices; so $E_i^p\cdot E_q^j$ is an integral chain of K^r.

To verify that $E_i^p\cdot E_q^j$ is an intersection, it is sufficient, in view of (1), (2), to prove the formula III (p. 350, loc. cit.):

III. $F(E_i^p\cdot E_q^j)=(FE_i^p)\cdot E_q^j+(-1)^p E_i^p\cdot(FE_q^j)$.

When $q=p-1$, $r=0$, (2) gives $E_i^p\cdot E_{p-1}^j=\mu_{ij}^p\sigma_i^p\sigma_j^{p-1}$. Using the rule for the boundary of a simplex given by its vertices, we obtain $F(E_i^p\cdot E_{p-1}^j)=\mu_{ij}^p(\sigma_j^{p-1}-\sigma_i^p)$. By formula (2) (loc. cit.), its dual, and (1),

$$(FE_i^p)\cdot E_{p-1}^j=\sum_k \mu_{ik}^p E_k^{p-1}\cdot E_{p-1}^j=\mu_{ij}^p\sigma_j^{p-1},$$

$$E_i^p\cdot(FE_{p-1}^j)=(-1)^{p+1}\mu_{ij}^p\sigma_i^p$$

which verifies III in this case. When $q<p-1$,

$$\sum_k \mu_{ik_1}^p\cdots\mu_{k_{s-1}k_s}^{p-s+1\ p-s}\mu_{k_sk_{s+1}}\cdots\mu_{k_rj}^{q+1\ p}\sigma_i\sigma_{k_1}^{p-1}\cdots\sigma_{k_{s-1}}^{p-s+1}\sigma_{k_{s+1}}^{p-s-1}\cdots\sigma_{k_r}^{q+1}\sigma_j^q=0,\qquad 1\leqq s\leqq r,$$

since $FF=0$ implies $\sum_k \mu_{k_{s-1}k_s}^{p-s+1}\mu_{k_sk_{s+1}}^{p-s}=0$. Hence

$$\begin{aligned}F(E_i^p\cdot E_q^j)&=\sum_k \mu_{ik_1}^p\cdots\mu_{k_rj}^{q+1}(\sigma_{k_1}^{p-1}\cdots\sigma_{k_r}^{q+1}\sigma_j^q\\&\quad+(-1)^{p-q}\sigma_i^p\sigma_{k_1}^{p-1}\cdots\sigma_{k_r}^{q+1})\\&=\sum_{k_1}\mu_{ik_1}^p(E_{k_1}^{p-1}\cdot E_q^j)+(-1)^{p-q}\sum_{k_r}\mu_{k_rj}^{q+1}(E_i^p\cdot E_{q+1}^{k_r})\end{aligned}$$

which by Lefschetz's article, formulas (2) and (1), is the right side of III.

CORNELL UNIVERSITY AND
INSTITUTE FOR ADVANCED STUDY

ON TRANSLATIONS OF FUNCTIONS AND SETS[1]

RALPH PALMER AGNEW

1. **Introduction.** It is the object of this note to prove the following theorem and two lemmas (see §3) on translations of sets which are used in the proof of the theorem.

THEOREM 1. *In order that a sequence $x_n(t)$ of complex-valued functions measurable over $-\infty<t<\infty$ may be such that, for each real sequence λ_n,*

$$\lim_{n\to\infty} x_n(t-\lambda_n)=0 \tag{1}$$

for almost all t, it is necessary and sufficient that for each $\delta>0$

$$\sum_{n=1}^{\infty} \text{l.u.b.}_{-\infty<h<\infty} \left| E_t\{h\leqq t\leqq h+1;\ |x_n(t)|\geqq\delta\}\right|<\infty. \tag{2}$$

Necessity for Theorem 1 is established by proving the following more incisive theorem.

THEOREM 2. *If a sequence $x_n(t)$ of complex-valued functions measurable over $-\infty<t<\infty$ is such that, for each real sequence λ_n,*

$$\lim_{n\to\infty} x_n(t-\lambda_n)=0$$

for each t in some set D of positive measure (where the set D may depend upon the sequence λ_n), then (2) holds.

Measure is that of Lebesgue, and a property such as (1) holds for almost all t if it holds for all t in the infinite interval $-\infty<t<\infty$ with the possible exception of a null set (set of measure 0). The set

$$A\equiv A(h,t,n,\delta)=E_t\{h\leqq t\leqq h+1;\ |x_n(t)|\geqq\delta\}$$

is the set of all points t such that $h\leqq t\leqq h+1$ and $|x_n(t)|\geqq\delta$; and $|A|$ denotes the measure of A. The condition (2) implies that when n is large the function $|x_n(t)|$ is less than δ for "most" values of t in each unit interval; but (2) implies no restriction whatever on $x_n(t)$ when t lies in the "exceptional" set.

The hypothesis that (1) holds for almost all t for each real *bounded* sequence λ_n does not imply (2). For example if, for each $n=1, 2, 3, \cdots$, $x_n(t)$ is a constant c_n over the interval $2^n<t<2^n+1$ and is 0 otherwise, and λ_n is a bounded sequence, then (1) holds for

[1] Presented to the Society, September 8, 1939.

each t; but (2) fails in case c_n fails to converge to 0 as n becomes infinite.

2. **Proof of sufficiency for Theorem 1.** Let $x_n(t)$ be a sequence of measurable functions for which (2) holds, and let λ_n be a sequence of real numbers. It follows from (2) that, for each $\delta>0$,

$$(3) \qquad \sum_{n=1}^{\infty} \text{l.u.b.}_{-\infty<h<\infty} \left| E_t\{h \leqq t \leqq h+1;\ |x_n(t-\lambda_n)| \geqq \delta\} \right| < \infty .$$

Let J denote an arbitrary finite interval. Since J can be covered by a finite set of unit intervals $h \leqq t \leqq h+1$, it follows from (3) that for each $\delta>0$

$$(4) \qquad \sum_{n=1}^{\infty} \left| E_t\{t \in J;\ |x_n(t-\lambda_n)| \geqq \delta\} \right| < \infty .$$

Setting

$$(5) \qquad A_{n,p} = E_t\{t \in J;\ |x_n(t-\lambda_n)| \geqq p^{-1}\}, \qquad n, p = 1, 2, 3, \cdots,$$

we see that (4) implies existence of indices $n_1<n_2<n_3<\cdots$ such that

$$(6) \qquad \sum_{n=n_p}^{\infty} |A_{n,p}| < 2^{-p-1}, \qquad p = 1, 2, \cdots .$$

Setting

$$(7) \qquad A_r = \sum_{p=r}^{\infty} \sum_{n=n_p}^{\infty} A_{n,p}, \qquad r = 1, 2, \cdots,$$

we find

$$(8) \qquad |A_r| \leqq \sum_{p=r}^{\infty} \sum_{n=n_p}^{\infty} |A_{n,p}| < \sum_{p=r}^{\infty} 2^{-p-1} = 2^{-r}, \qquad r = 1, 2, \cdots .$$

Let

$$(9) \qquad J_r = J - A_r, \qquad r = 1, 2, \cdots .$$

If $t \in J_r$ then, when $p>r$,

$$(10) \qquad |x_n(t-\lambda_n)| < p^{-1}, \qquad n \geqq n_p,$$

so that $x_n(t-\lambda_n)$ converges to 0 over J_r. Hence $x_n(t-\lambda_n)$ converges to 0 over $J_1+J_2+\cdots$. But J_r is a subset of J having measure greater than $|J|-2^{-r}$; hence $J_1+J_2+\cdots$ is a subset of J having measure $|J|$. Therefore $x_n(t-\lambda_n)$ converges to 0 for almost all t in J.

Since J is an arbitrary finite interval, $x_n(t-\lambda_n)$ must converge to 0 for almost all t in $-\infty<t<\infty$ and sufficiency for Theorem 1 is proved.

3. **Lemmas on translations of sets.** In this section we prove two lemmas. The first states that if C and B are measurable subsets of unit intervals, then it is possible to translate B in such a way that the intersection of C and the translation of B will have measure at least $\frac{1}{2}|C||B|$. The first lemma is used in proof of the second which specifies conditions under which a given sequence of sets can be translated so as to cover each point of the interval $-\infty<t<\infty$, with the exception of a null set, an infinite number of times. The close connection established in §4 between Lemma 2 and Theorem 2 shows that the combined proofs of Lemmas 1 and 2 furnish essentially a proof of Theorem 2.

If E is a set of points t in the interval $-\infty<t<\infty$ and λ is a real number, let $E(\lambda)$ denote the set of points t such that $t-\lambda \in E$; thus $E(\lambda)$ is the set obtained by translating the set E to the right λ units. Let U denote the unit interval $0\leqq t\leqq 1$.

LEMMA 1. *If C and B are measurable subsets of U, then*

$$\max_{-1\leqq\lambda\leqq 1} |CB(\lambda)| \geqq \tfrac{1}{2}|C||B|. \tag{11}$$

Let $\phi(t)$ be the characteristic function of C, that is, $\phi(t)=1$ when $t \in C$ and $\phi(t)=0$ otherwise; and let $\psi(t)$ be the characteristic function of B. Then $\psi(t-\lambda)$ is the characteristic function of $B(\lambda)$, and $\phi(t)\psi(t-\lambda)$ is the characteristic function of the intersection $CB(\lambda)$ of C and $B(\lambda)$. Hence on denoting the measure of $CB(\lambda)$ by $\mu(\lambda)$ we have

$$\mu(\lambda) = \int_{-\infty}^{\infty} \phi(t)\psi(t-\lambda)dt. \tag{12}$$

The function $\mu(\lambda)$ is continuous since

$$\begin{aligned} |\mu(\lambda+h)-\mu(\lambda)| &\leqq \int_{-\infty}^{\infty} \phi(t)\,|\psi(t-\lambda-h)-\psi(t-\lambda)|\,dt \\ &\leqq \int_{-\infty}^{\infty} |\psi(t-\lambda-h)-\psi(t-\lambda)|\,dt \\ &= \int_{-\infty}^{\infty} |\psi(t-h)-\psi(t)|\,dt \end{aligned}$$

and the last integral converges to 0 with h. Hence $\mu(\lambda)$ has a maxi-

mum over the interval $-1 \leqq \lambda \leqq 1$. Since $\mu(\lambda) = 0$ when $|\lambda| > 1$, the computation

$$\int_{-1}^{1} \mu(\lambda)\,d\lambda = \int_{-\infty}^{\infty} d\lambda \int_{-\infty}^{\infty} \phi(t)\psi(t-\lambda)\,dt$$

$$= \int_{-\infty}^{\infty} \phi(t)\,dt \int_{-\infty}^{\infty} \psi(t-\lambda)\,d\lambda = |C|\,|B|$$

is easily justified. This equality and the inequality

$$\mu(\lambda) \leqq \max_{-1 \leqq \lambda \leqq 1} \mu(\lambda), \qquad -1 \leqq \lambda \leqq 1, \tag{13}$$

imply that $\max |CB(\lambda)| = \max \mu(\lambda) \geqq \frac{1}{2}|C|\,|B|$ and Lemma 1 is established.

The fact that use of inequalities such as (13) often leads to crude results may make one suspicious that Lemma 1 holds when the factor $\frac{1}{2}$ in (11) is replaced by a greater factor. To settle this question, let $0 < \epsilon < \frac{1}{3}$, let $C = E_t\{\epsilon \leqq t \leqq 1-\epsilon\}$, and let $B = E_t\{0 \leqq t \leqq \epsilon\} + E_t\{1-\epsilon \leqq t \leqq 1\}$. Then $|C| = 1-2\epsilon$, $|B| = 2\epsilon$, and it is easy to verify that

$$\max_{-1 \leqq \lambda \leqq 1} |CB(\lambda)| = \epsilon = [1/(2-4\epsilon)]\,|C|\,|B| > 0. \tag{14}$$

This shows that $\frac{1}{2}$ is the greatest factor permissible in (11).

LEMMA 2. *If $A_1, A_2, \cdots$ is a sequence of measurable sets and a sequence $U_1, U_2, \cdots$ of unit intervals exists such that*

$$\sum_{n=1}^{\infty} |U_n A_n| = \infty, \tag{15}$$

then there exists a sequence $\lambda_1, \lambda_2, \cdots$ such that each t in the interval $-\infty < t < \infty$, except those in some null set, lies in an infinite number of the sets $A_n(\lambda_n)$.

Let $B_n = U_n A_n$ so that each B_n lies in some unit interval and $\sum |B_n| = \infty$. Let n be fixed. Choose λ_n such that $B_n(\lambda_n) \subset U$, where U is as before the unit interval $0 \leqq t \leqq 1$, and let $C_n = U - B_n(\lambda_n)$. Since λ'_{n+1} exists such that $B_{n+1}(\lambda'_{n+1}) \subset U$, Lemma 1 guarantees existence of λ_{n+1} such that

$$|C_n B_{n+1}(\lambda_{n+1})| \geqq \tfrac{1}{2}|C_n|\,|B_{n+1}|. \tag{16}$$

Let $C_{n+1} = U - [B_n(\lambda_n) + UB_{n+1}(\lambda_{n+1})]$. Again from Lemma 1, λ_{n+2} exists such that (16) holds when n is replaced by $n+1$. In this manner,

we obtain a sequence $\lambda_n, \lambda_{n+1}, \cdots$ of real numbers and a sequence

$$C_{n+p} = U - [UB_n(\lambda_n) + UB_{n+1}(\lambda_{n+1}) + \cdots + UB_{n+p}(\lambda_{n+p})] \tag{17}$$

of sets such that, for each $p=0, 1, 2, \cdots$,

$$\sum_{k=n}^{n+p} |C_k B_{k+1}(\lambda_{k+1})| \geqq \frac{1}{2} \sum_{k=n}^{n+p} |C_k|\,|B_{k+1}|. \tag{18}$$

Since the sets $C_k B_{k+1}(\lambda_{k+1})$ $(k=n, n+1, \cdots, n+p)$ are subsets of U and no two have a point in common, the left member of (18) is less than or equal to unity for each $p=0, 1, 2, \cdots$. From this it follows that $|C_{n+p}| \to 0$ as $p \to \infty$; for $|C_{n+p}|$ is monotone decreasing as $p \to \infty$ and if $|C_{n+p}|$ is bounded from 0, then the fact that $\sum |B_n| = \infty$ would imply that the right member of (18) diverges to $+\infty$ as $p \to \infty$. The conclusion that $|C_{n+p}| \to 0$ as $p \to \infty$ implies by (17) that

$$\lim_{p \to \infty} |UB_n(\lambda_n) + UB_{n+1}(\lambda_{n+1}) + \cdots + UB_{n+p}(\lambda_{n+p})| = 1. \tag{19}$$

Hence there exists a sequence $0=n_1<n_2<\cdots$ of indices such that the set

$$D_k \equiv UB_{n_k+1}(\lambda_{n_k+1}) + \cdots + UB_{n_{k+1}}(\lambda_{n_{k+1}}) \tag{20}$$

has measure $|D_k| > 1-2^{-k-1}$ for each $k=1, 2, \cdots$. Put $P_k = D_k D_{k+1} \cdots$ and $P = P_1 + P_2 + \cdots$. The fact that $D_k \subset U$ and $|D_k| > 1-2^{-k-1}$ for each $k=1, 2, \cdots$ implies that $P_k \subset U$ and $|P_k| \geqq 1-2^{-k}$, and consequently $P \subset U$ and $|P| = 1$. If $t \in P$, then $t \in P_k$ for some k so that $t \in D_k$ for all sufficiently great k and $t \in B_n(\lambda_n)$ for an infinite set of n, and hence also $t \in A_n(\lambda_n)$ for an infinite set of n.

If the sequence of sets A_n is arranged in a double sequence $A_{p,q}$ $(p=0, \pm 1, \cdots; q=1, 2, \cdots)$ in such a way that

$$\sum_{q=1}^{\infty} |A_{p,q}| = \infty, \quad p = 0, \pm 1, \pm 2, \cdots, \tag{21}$$

it results from what we have already proved that for each fixed p there is a sequence $\lambda_{p,1}, \lambda_{p,2}, \cdots$ such that each point of a subset of $I_p \equiv E_t\{p \leqq t \leqq p+1\}$ of measure unity is contained in an infinite number of the sets $A_{p,1}(\lambda_{p,1}), A_{p,2}(\lambda_{p,2}), \cdots$. Then each point of $-\infty < t < \infty$ with the exception of a null set lies in an infinite number of sets of the double sequence $A_{p,q}(\lambda_{p,q})$ which can be arranged in the simple sequence $A_n(\lambda_n)$, and proof of Lemma 2 is complete.

The hypothesis of Lemma 2 is equivalent to the following: A_n is a sequence of measurable sets such that

$$(22) \qquad \sum_{n=1}^{\infty} \text{l.u.b.}_{-\infty<h<\infty} \left| E_t\{h \leqq t \leqq h+1; t \in A_n\} \right| = \infty .$$

That the hypothesis (22) cannot be relaxed is a consequence of the following result which we give without proof. If $A_1, A_2, \cdots$ is a sequence of measurable sets, and a real sequence $\lambda_1, \lambda_2, \cdots$ and a set C of positive measure exist such that each point of C lies in an infinite number of the sets $A_n(\lambda_n)$, then (22) holds.

That the conclusion of Lemma 2 must provide for an exceptional null set becomes clear when one observes that if the sets A_n are each nondense then, however $\lambda_1, \lambda_2, \cdots$ are determined, the set $\sum A_n(\lambda_n)$ must be of the first category and hence there must be a set of the second category whose points are in *none* of the sets $A_n(\lambda_n)$.

4. **Proof of Theorem** 2. To prove Theorem 2, let $x_n(t)$ be a sequence of measurable functions for which (2) fails for some $\delta > 0$. Then $\delta > 0$ and a sequence $h_1, h_2, \cdots$ exist such that

$$(23) \qquad \sum_{n=1}^{\infty} E_t\{h_n \leqq t \leqq h_n + 1; \left| x_n(t) \right| \geqq \delta\} = \infty .$$

Let $A_n = E_t\{\left| x_n(t) \right| \geqq \delta\}$. Then by Lemma 2 there exist a sequence $\lambda_1, \lambda_2, \cdots$ and a set C whose complement is a null set such that each t in C lies in an infinite number of the sets $A_n(\lambda_n)$. Hence if $t \in C$, then $t - \lambda_n \in A_n$ for an infinite set of n so that $\left| x_n(t-\lambda_n) \right| \geqq \delta$ for an infinite set of n. This contradicts the hypothesis of Theorem 2 and completes the proof.

CORNELL UNIVERSITY

A NEW FORMULA FOR THE BERNOULLI NUMBERS[1]

H. L. GARABEDIAN

It is the purpose of this note to exhibit a unique method for deriving what appears to be a new formula for the Bernoulli numbers. We obtain the formula

$$B_{k+1} = \frac{(-1)^{k+1}(k+1)}{2^{k+1}-1} \sum_{i=0}^{k} \frac{\Delta^i a_0}{2^{i+1}}, \tag{1}$$

$$a_n = (1+n)^k, \quad k = 0, 1, 2, \cdots,$$

or

$$B_{k+1} = \frac{(-1)^{k+1}(k+1)}{2^{k+1}-1}\left[\frac{1}{2} + \frac{1}{2^2}(1-2^k) + \frac{1}{2^3}(1-2\cdot 2^k+3^k) + \cdots + \frac{1}{2^{k+1}}\{1 - C_{k,1}2^k + C_{k,2}3^k - + \cdots + (-1)^k C_{k,k}(k+1)^k\}\right],$$

$$k = 0, 1, 2, \cdots.$$

Some of the Bernoulli numbers computed from (1) are $B_1 = -1/2$, $B_2 = 1/6$, $B_3 = 0$, $B_4 = -1/30$, $B_5 = 0$, $B_6 = 1/42$, $B_7 = 0$, $B_8 = -1/30$, $B_9 = 0$, $B_{10} = 5/66$. Evidently, the known formulas[2] which give the Bernoulli numbers in explicit form, in comparison with formula (1), yield only the numerical values of these numbers with even indices. While none of the formulas under discussion could possibly serve a useful purpose in computing unknown numbers with high indices, formula (1) appears to be the simplest in form and the one best adapted to computation of numbers with low indices.

Our method of obtaining (1) consists in summing the divergent series

$$\sum_{n=0}^{\infty} (-1)^n (n+1)^k = 1 - 2^k + 3^k - + \cdots, \quad k = 0, 1, 2, \cdots, \tag{2}$$

using two consistent methods of summation, one of which assigns a value to (2) which involves the Bernoulli numbers. Equating the two values thus obtained we get the desired formula.

First, we sum the series (2) by the method of Abel. To this end we

---

[1] Presented to the Society, December 29, 1939.

[2] See, for example, Niels Nielsen, *Traité Élémentaire des Nombres de Bernoulli*, 1923.

have to make some preliminary observations. The Bernoulli numbers may be defined by the relation[3]

$$\frac{x}{e^x - 1} = B_0 + \frac{B_1}{1!}x + \frac{B_2}{2!}x^2 + \cdots, \tag{3}$$

which is valid at least for small values of x. Now, consider the series

$$f(t) = e^{-t} - e^{-2t} + e^{-3t} - + \cdots,$$

which converges for $e^{-t} < 1$ and hence for $t > 0$. Moreover, we can write

$$\begin{aligned} f(t) &= \frac{e^{-t}}{1 + e^{-t}} = \frac{1}{e^t + 1} = \frac{e^t + 1 - 2}{e^{2t} - 1} \\ &= \frac{1}{e^t - 1} - \frac{2}{e^{2t} - 1} = \frac{1}{t}\,\frac{t}{e^t - 1} - \frac{1}{t}\,\frac{2t}{e^{2t} - 1}. \end{aligned} \tag{4}$$

We may now expand the last expression in (4) using (3). We obtain for small values of $t > 0$ the relation

$$\begin{aligned} f(t) = \frac{1}{t}\left[B_0 + \frac{B_1}{1!}t + \frac{B_2}{2!}t^2 + \cdots\right] \\ - \frac{1}{t}\left[B_0 + \frac{2B_1}{1!}t + \frac{2^2B_2}{2!}t^2 + \cdots\right], \end{aligned}$$

or

$$\sum_{n=0}^{\infty} (-1)^n e^{-(n+1)t} = -\sum_{n=0}^{\infty} \frac{2^{n+1} - 1}{(n+1)!} B_{n+1} t^n.$$

Differentiating k times with respect to t, we have

$$\begin{aligned} &\sum_{n=0}^{\infty} (-1)^n (n+1)^k e^{-(n+1)t} \\ &\qquad = (-1)^{k+1} \sum_{n=k}^{\infty} \frac{2^{n+1} - 1}{(n+1)!} B_{n+1} \cdot n(n-1) \cdots (n-k+1) t^{n-k}. \end{aligned}$$

Now, allowing t to approach zero through positive values, we obtain

$$\lim_{t \to 0+} \sum_{n=0}^{\infty} (-1)^n (n+1)^k e^{-(n+1)t} = (-1)^{k+1} \frac{2^{k+1} - 1}{k+1} B_{k+1}. \tag{5}$$

Putting $x = e^{-t}$ in the series on the left in (5), we have a power series

[3] Euler, *Institutiones Calculi Differentialis* . . . , vol. 2, 1755, §122.

with radius 1. Thus, the statement (5) implies that the series (2) is summable by the method of Abel (or is summable A) to the value given in (5).[4]

Next, we sum the series (2) by the Cesàro method. We need the following theorem.[5]

THEOREM. *The series* $\sum_{n=0}^{\infty}(-1)^n a_n$ *is exactly summable* (C, ν), $(\nu = 1, 2, 3, \cdots)$, *to the value*

$$\sum_{n=0}^{\nu-1} \frac{\Delta^n a_0}{2^{n+1}}$$

provided that

$$\Delta^i a_0 = 0, \qquad \Delta^{\nu-1} a_0 \neq 0, \qquad i \geqq \nu \geqq 1.$$

Since the general term in the series (2) is a polynomial of degree k, its differences of order higher than k all vanish. Thus, due to our theorem, the series (2) is summable $(C, k+1)$ to the value

$$\sum_{n=0}^{k} \frac{\Delta^n a_0}{2^{n+1}}, \qquad a_n = (n+1)^k, \; k = 0, 1, 2, \cdots. \tag{6}$$

Now, it is well known that Abel summability includes Cesàro summability of positive integral orders. Then, the series (2) must be summable A to the value (6). Equating (5) and (6) we get the desired formula:

$$B_{k+1} = \frac{(-1)^{k+1}(k+1)}{2^{k+1}-1} \sum_{i=0}^{k} \frac{\Delta^i a_0}{2^{i+1}}, \qquad a_n = (n+1)^k, \; k = 0, 1, 2, \cdots.$$

NORTHWESTERN UNIVERSITY

[4] K. Knopp, *Theorie und Anwendung der unendlichen Reihen*, 1922, 2d edition, 1924, p. 508.

[5] H. L. Garabedian, this Bulletin, vol. 45 (1939), pp. 592–596.

ON A PROBLEM CONCERNING PROBABILITY AND ITS CONNECTION WITH THE THEORY OF DIFFUSION[1]

M. KAC

Introduction. The present note gives a certain probabilistic approach to the problem of diffusion. The main result is that the classical solution of the differential equation of diffusion is an asymptotic formula for the statistical problem under consideration.

The method which will be used is essentially that of Steinhaus and the present author which they applied to a similar but simpler problem of P. and T. Ehrenfest.[2]

1. **The problem.** Given an infinite sequence of boxes enumerated as follows

$$\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots$$

and N numbered balls which are distributed in a certain way in the boxes, one takes at random one of the numbers $-N, \cdots, -1, 1, \cdots, N$ which are suppose to be equiprobable and if the number k is drawn one moves the ball number $|k|$ from its original box to the nearest to the right or to the nearest to the left according as sign k was 1 or -1. One repeats this process n times and one asks what is the "probable value" (mathematical expectation) of the "concentration" of balls in the box number s, say. By "concentration" one simply understands the ratio of the number of balls in a certain box and N. It is, of course, understood that the successive drawings are independent in the statistical sense of this word.

2. **Reduction of the problem by means of the "function of choice."** We divide the interval (0, 1) into N equal parts and we define a function on (0, 1) by placing $f(x)=s$ for $l-1/N<x\leqq l/N$ if the ball number l is originally in box s. This function represents the initial state of the schema. The joint length of those intervals in which $f(x)=s$ is obviously the initial "concentration" of balls in the box number s.

Let now $\omega(x)$ (the "function of choice") be 1 for $0<x\leqq 1/N$ and 0 for $1/N<x\leqq 1$ and let furthermore $\omega(x+1)=\omega(x)$. Then $f(x)\pm\omega(x-(p-1)/N)$ represents obviously the state of the schema after moving the ball number p from its box to the nearest to the right $(+)$ or to the nearest to the left $(-)$.

[1] Presented to the Society, December 29, 1939.

[2] H. Steinhaus, *La Théorie et les Applications des Fonctions Indépendantes*, Actualités Scientifiques et Industrielles, Paris, 1938.

Let r be a "random variable" which takes only the values -1, $-(N-1)/N, \cdots, -1/N, 1/N, \cdots, (N-1)/N, 1$ and every one of them with the probability $1/2N$. Then $f(x)+\text{sign } r\cdot\omega(x-|r|+1/N)$ represents the "random state" after the first drawing. If $r_1, r_2, \cdots, r_n$ are independent "random variables" having the same distribution as r, then

$$\Sigma_n(x, r) = f(x) + \text{sign } r_1\omega(x - |r_1| + 1/N) + \cdots + \text{sign } r_n\omega(x - |r_n| + 1/N)$$

represents the "random state" after n independent drawings. Thus the problem is reduced to evaluating the "probable value" of the joint length of those intervals in which $\Sigma_n(x, r)=s$.

3. **Evaluation by means of a "discontinuity factor."** The remaining part of the solution is now of a purely technical nature. It is clear that

$$\frac{1}{2\pi}\int_0^{2\pi} \exp\,[i\xi(\Sigma_n(x, r) - s)]d\xi$$

is equal to 1 or to 0 according as $\Sigma=s$ or $\Sigma\neq s$, and therefore the joint length of those intervals in which $\Sigma=s$ is

$$L_n(r) = \frac{1}{2\pi}\int_0^1\int_0^{2\pi} \exp\,[i\xi(\Sigma_n(x, r) - s)]d\xi dx.$$

Thus the "probable value" $P_{n,N}(s)$ of $L_n(r)$ is equal to

$$\frac{1}{2\pi}\int_0^1\int_0^{2\pi} \text{prob. val. } \{\exp[i\xi(\Sigma_n(x, r) - s)]\}d\xi dx.$$

It is easily seen that

$$\begin{aligned}&\text{prob. val. } \{\exp\,[i\xi(\Sigma - s)]\}\\ &= \exp\,[i\xi(f(x) - s)](\text{prob. val. } \exp\,[(i\xi \text{ sign } r)\omega(x - |r| + 1/N)])^n\\ &= \exp\,[i\xi(f(x) - s)]\left(1 - \frac{1}{N} + \frac{1}{N}\cos\xi\right)^n\end{aligned}$$

and finally

$$(1)\quad P_{n,N}(s) = \frac{1}{2\pi}\int_0^1\int_0^{2\pi} \cos\xi(f(x) - s)\left(1 - \frac{1}{N} + \frac{1}{N}\cos\xi\right)^n d\xi dx.$$

4. **A particular case.** The formula becomes much simpler if one assumes that $f(x)\equiv 0$, that is, all the balls were initially in the box number 0. In this case one has

$$P_{n,N}(s) = \frac{1}{2\pi}\int_0^{2\pi} \cos s\xi\left(1 - \frac{1}{N} + \frac{1}{N}\cos\xi\right)^n d\xi.$$

Suppose that $N\to\infty$ and that $n/N\to z$; then

$$P_{n,N}(s) \to \frac{e^{-z}}{2\pi}\int_0^{2\pi} \cos s\xi e^{z\cos\xi} d\xi = e^{-z}J_s(iz),$$

where $J_s(z)$ denotes the sth Bessel function.

5. **The asymptotic formula.** Suppose now that the boxes are situated on the infinite line $(-\infty, +\infty)$ and that the distance between the consecutive ones is λ. Suppose furthermore that one makes a drawing every τ seconds so that after the time t one has performed $n = [t/\tau]$ drawings. We shall be interested in evaluating the "probable value" of the "concentration" of balls on the interval (l_1, l_2) after the time t. The answer is given by the probable value of the joint length of those intervals in which $l_1 < \lambda\Sigma_n(x, \tau) < l_2$ and can be evaluated by means of the formula (1). However, it will be more convenient for our purpose to evaluate the "probable value" under consideration by means of the Dirichlet discontinuity factor. It is well known that

$$\frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{\sin\alpha\xi}{\xi} e^{i\gamma\xi} d\xi$$

is equal to 1, $\frac{1}{2}$ or 0 according as $|\gamma| < \alpha$, $|\gamma| = \alpha$ or $|\gamma| > \alpha$. Thus, if neither l_1 nor l_2 is a multiple of λ, the probable value one looks for is given by the following formula:

$$P_N(l_1, l_2, t) = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{\sin\frac{1}{2}(l_2 - l_1)\xi}{\xi} \cdot\exp\left(-i\xi\frac{l_1 + l_2}{2}\right)\left(\int_0^1 \exp\left(i\xi\lambda f(x)\right)dx\right) \cdot\left(1 - \frac{1}{N} + \frac{1}{N}\cos\lambda\xi\right)^{[t/\tau]} d\xi.$$

If now $\lambda^2/N\tau \to 2k$ as $N\to\infty$, and $\lambda\to 0$, one has

$$\left(1 - \frac{1}{N} + \frac{1}{N}\cos\lambda\xi\right)^{[t/\tau]} \to \exp\left(-kt\xi^2\right)$$

uniformly in every finite interval and, assuming in addition that the distribution function of $\lambda f(x)$ tends to a distribution function $\sigma(u)$,

one readily deduces from the continuity theorem of Fourier-Stieltjes transforms[3] that

$$P_N(l_1, l_2, t) \to \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{\sin \frac{1}{2}(l_2 - l_1)\xi}{\xi} \cdot \exp\left(-i\xi\,\frac{l_1 + l_2}{2}\right)\left(\int_{-\infty}^{+\infty} \exp\,(i\xi u)d\sigma(u)\right)\exp\,(-\,kt\xi^2)d\xi.$$

The integral on the right side can easily be evaluated and its value is

$$\frac{1}{2(\pi kt)^{1/2}}\int_{l_1}^{l_2}\int_{-\infty}^{+\infty} \exp\left\{-\frac{(u-l)^2}{4kt}\right\}d\sigma(u)dt.$$

In the limiting case the concentration is given by the formula

$$\frac{1}{2(\pi kt)^{1/2}}\int_{-\infty}^{+\infty} \exp\left\{-\frac{(u-l)^2}{4kt}\right\}d\sigma(u),$$

which is the classical solution of the differential equation of diffusion in the case of an infinite cylinder. In our case the classical formula turned out to be an asymptotic formula for a simple probabilistic process belonging to the category of the so-called "random walks."

CORNELL UNIVERSITY

[3] Cf. E. K. Haviland, American Journal of Mathematics, vol. 56 (1934), pp. 625–658.

THE SOLUTION OF HARMONIC EQUATIONS BY MEANS OF DEFINITE INTEGRALS

H. BATEMAN

Darboux devotes a chapter of his famous *Théorie des Surfaces* to the study of partial differential equations related to some equations of mathematical physics that have sets of simple solutions of type $X(x)\,Y(y)$. Such equations are called harmonic equations and are of great interest on account of the various ways in which they can be reduced to the harmonic form and solved by definite integrals. The reduction problem has already been more or less solved, but it cannot be said that the definite integral problem has been fully discussed.

It is known, for instance, that with certain special forms of the functions $a(\theta)$, $b(\theta)$, $h(\theta)$, $f(\theta)$ an integral of type

$$z = \int_0^\pi F(ax^2 + 2hxy + by^2) f(\theta)\, d\theta \tag{1}$$

may satisfy the harmonic equation

$$r + t + up/x + vq/y = 0, \qquad u = 2k + 1,\ v = 2m + 1, \tag{2}$$

where u and v are constants and p, q, r, s, t denote the partial derivatives of z of the first two orders. For general values of u and v this equation is of interest in the study of solutions of the wave equation and Laplace's equation in four variables. When u and v are integers, the equation arises in the study of symmetrical solutions of Laplace's equation in N variables, in the study of the stream function of hydrodynamics and of various functions which occur in the theory of elasticity.

To find all the possible forms of $a(\theta)$, $b(\theta)$, $h(\theta)$, $f(\theta)$ when it is assumed that the integral can be differentiated in the usual way, we must discover how to satisfy the equation

$$\begin{aligned} 0 = \int_0^\pi f(\theta)\, d\theta \big[& 4F''(w)\{(ax + by)^2 + (hx + by)^2\} \\ & + 2F'(w)\{a + b + u(a + hy/x) + v(b + hx/y)\}\big] \end{aligned} \tag{3}$$

where $w = ax^2 + 2hxy + by^2$. As there are many different ways in which this might perhaps be done, it is not yet known whether all the useful types of integral have already been found.

If the limits were $-\pi$ and π, we could satisfy the equation by making the integrand the sum of an exact differential and an odd function

of θ. When the limits are 0 and π, there is the possibility that part of the integral may be zero on account of an invariance in form but with sign changed when θ is replaced by $\pi-\theta$. There are other possibilities suggested by the theory of functional differential equations, but at present we shall consider only the case in which the equation is satisfied because the integrand is equal to $d[GF'(w)]$ where G depends only on x, y and θ in such a way that $GF'(w)$ has the same finite value at both limits. In this case the analysis depends only on the ordinary theory of exact differentials, and it is not necessary to enter into extensions of this idea and extensions of the theory of the adjoint differential equation. Such extensions are occasionally useful but are not hard to develop.

Using primes to denote derivatives of a, b, h, f, w and G with respect to θ, the equations to be satisfied are

$$(4)\qquad \begin{aligned} &Gw' = 4f(A^2 + B^2), \qquad A = ax + hy,\ B = hx + by,\\ &xyG' = 2f[(a+b)xy + uyA + vxB]. \end{aligned}$$

Eliminating G we obtain an equation which can be satisfied only if $uhb'=0$ and $vha'=0$. The three possibilities are (1) $u=0$, $a'=0$, (2) $v=0$, $b'=0$, (3) $h=0$.

In the first case a is constant and the equations become

$$(4')\qquad yG' = 2f[(a+b)y + vB], \qquad Gy = 4f(A^2 + B^2)/w_0$$

where $w_0=2h'x+b'y$. Eliminating yG we obtain the equation

$$(5)\qquad \begin{aligned} &4w_0(h'yA + BB' + \tfrac{1}{2}f'A^2/f + \tfrac{1}{2}f'B^2/f)\\ &\qquad = 2w_0'(A^2 + B_0^2) + w^2[(a+b)y + vB], \end{aligned}$$

which should hold for all values of x and y. Choosing the ratio of these quantities in such a way that $w_0=0$, we obtain the equation

$$(6)\qquad (h''b' - b''h')[(ab' - 2hh')^2 + (hb' - 2bh')^2] = 0.$$

The first factor vanishes when h'/b' is a constant. The second factor vanishes when $(a+ih)b'=2(h+ib)h'$ or

$$(7)\qquad b = 2(a + ih) - a + c(a + ih)^2,$$

where c is an arbitrary constant.

Equating the coefficients of x^3 in (5), we get

$$h''/h' - f'/f = (2 - v)hh'/(a^2 + h^2).$$

Hence $fd\theta = dh(a^2+h^2)^{(v-1)/2}$, the constant of integration being put to unity.

An appropriate value of the constant c may be found by choosing the ratio of x/y in (5) so that $A^2+B^2=0$ but $w_0 \neq 0$. The value $c=-1/a$ makes $w=A^2/a$ and the integral becomes the well known integral of Poisson[1] if we take $h=-i\cos\theta$.

If we try

$$z = \int_{-\pi}^{\pi} H(x - iy\cos\theta) f(\theta) d\theta, \tag{8}$$

the equation $r+t+vq/y=0$ will be satisfied if

$$0 = [H' f \sin\theta]_{-\pi}^{\pi} - \int_{-\pi}^{\pi} H'[f' \sin\theta - (fv - 1)\cos\theta] d\theta. \tag{9}$$

The integral may be made to vanish by choosing f so that $-f'\sin\theta + f(v-1)\cos\theta$ is an odd function of θ. In the case when this function is $\sin^{2n+1}\theta$, the equation for the determination of f is

$$f'(\theta)\sin\theta - (v-1)f(\theta)\cos\theta = \sin^{2n+1}\theta \tag{10}$$

and the appropriate solution is

$$f(\theta) = \sin^{v-1}\theta \int_0^{\theta} \sin^{2n-v+1}\theta \, d\theta. \tag{11}$$

If, for example, $v=1$ and $n=0$, we get $f(\theta)=\theta$ and the integral is zero. The device does not yield a new solution.

Passing now to the case $h=0$, the equations (4) become

$$(4'') \quad G(a'x^2 + b'y^2) = 4(a^2x^2 + b^2y^2) f, \; G' = [2(1+u)a + 2(1+v)b] f.$$

Elimination of G gives an equation

$$\begin{aligned} &4(aa'x^2 + bb'y^2)(a'x^2 + b'y^2) - 2(a^2x^2 + b^2y^2)(a''x^2 + b''y^2) \\ &\quad + 2(f'/f)(a^2x^2 + b^2y^2)(a'x^2 + b'y^2) \\ &\qquad = [(1+u)a + (1+v)b](a'x^2 + b'y^2)^2, \end{aligned} \tag{12}$$

which must be satisfied for all values of x and y. Choosing the ratio of x and y so that $a'x^2+b'y^2=0$, we find that

$$(a''b' - a'b'')(a^2b' - b^2a') = 0. \tag{13}$$

The first factor vanishes when b'/a' is constant, the second factor when $b'/b^2=a'/a^2$ or $b^{-1}=a^{-1}+\text{const.}$ Taking the constant to be unity, we find that if $a=\tan^2\frac{1}{2}\theta$, $b=\sin^2\frac{1}{2}\theta$. Equating coefficients of x^4

[1] Journal de l'École Polytechnique, vol. 12 (1821), pp. 215–248. See also R. Hoppe, Journal für die reine und angewandte Mathematik, vol. 58 (1861), pp. 369–373.

in (12), we find that

$$(14)\quad 4(a'/a) - 2(a''/a') + 2f'/f = (1+u)(a'/a) + (1+v)(ba'/a^2).$$

Now $ba'/a^2 = \cot \frac{1}{2}\theta$ and consequently $f = (\tan \frac{1}{2}\theta)^u (\sin \frac{1}{2}\theta)^{v-1}$. It may be verified that if $w = x^2 \tan^2 \frac{1}{2}\theta + y^2 \sin^2 \frac{1}{2}\theta = W^2$ and F is a suitable function, which may even be zero for $w > c^2$,

$$(15)\quad z = \int_0^\pi F(w) \tan^u \tfrac{1}{2}\theta \sin^{v-1} \tfrac{1}{2}\theta \, d\theta$$

is indeed a solution of $r+t+up/x+vq/y=0$ when $R(u+v)>0$. A particular case of this result is that if $x>0$, $R(m) > -\frac{5}{4}$, $R(k+m) > -1$,

$$(16)\quad \int_0^\pi W^{-k-m} J_{k+m}(W) \sin^{2m} \tfrac{1}{2}\theta \tan^{2k+1} \tfrac{1}{2}\theta \, d\theta = 2x^{-k} K_k(x) y^{-m} J_m(y).$$

This result may be verified by using Lommel's expansion

$$(17)\quad W^{-k-m} J_{k+m}(W) = \sum_{n=0}^{\infty} (-\tfrac{1}{2}y^2 S^2)^n / n! (xT)^{-k-m-n} J_{k+m+n}(xT),$$

in which $S = \sin \frac{1}{2}\theta$, $T = \tan \frac{1}{2}\theta$. Now by a well known formula

$$(18)\quad \int_0^\pi J_{k+m+n}(xT) S^{2m+2n} T^{k+1-m-n} d\theta = 2(\tfrac{1}{2}x)^{m+n} K_k(x) / \Gamma(m+n+1),$$
$$-1 < R(k+m+n) < R(2m+2n+\tfrac{3}{2}).$$

Hence, if $R(k+m) > -1$, $R(k-m) < \frac{3}{4}$, $R(m) > -\frac{5}{4}$, the series

$$\sum_{n=0}^{\infty} (-\tfrac{1}{2}y^2)^n / n! \int_0^\pi S^{2m+2n} T^{2k+1} (xT)^{-k-m-n} J_{k+m+n}(xT) d\theta$$
$$= 2x^{-k} K_k(x) \sum_{n=0}^{\infty} (-y^2)^n 2^{-m-2n} / n! \Gamma(m+n+1)$$
$$= 2x^{-k} K_k(x) y^{-m} J_m(y).$$

The formula (16) is obtained by changing the order of integration and summation.[2]

If in the integral (15) w is taken as the new independent variable, the integrand should satisfy the partial differential equation (2) for all

[2] This may be justified by forming estimates of the remainders in the two series. When $u=0$, $v=1$, the integral (15) is useful in potential theory and gives integral expressions for many symmetrical potentials. A list of these is being prepared. One such expression occurs in the paper of E. G. Gallop, *The distribution of electricity on the circular disc and spherical bowl*, Quarterly Journal of Mathematics, vol. 21 (1886), pp. 229–256.

values of the parameter w. If $r=(2wx^2+2x^2y^2-2y^2w+x^4+y^4+w^2)^{1/2}$, the result is that

$$z = r^{-1}x^{-2k}(r + w - x^2 - y^2)^{m+k}(r + w + x^2 - y^2)^{-m} \tag{19}$$

should be a solution of the partial differential equation. The solution of the partial differential equation (2) may be written in a more symmetrical form by putting

$$w = \frac{x^2}{a^2 + v} + \frac{y^2}{b^2 + v}, \quad z = \int_C F(w)(a^2 + v)^{-k-1}(b + v)^{-m-1}dv.$$

The contour C may start at infinity and enclose either $v=-a^2$ or $v=-b^2$. It may also run from $-\infty$ to $+\infty$ along a path which avoids the points $v=-a^2$ and $v=-b^2$. The solution is now seen to be related to the integral for the product of two Bessel functions obtained by S. O. Rice, Quarterly Journal of Mathematics (Oxford series), vol. 6 (1935), pp. 52–64.

This theorem may be extended to any number of variables of type x and y. Thus if w is defined by the equation

$$w = \frac{x^2}{a^2 + v} + \frac{y^2}{b^2 + v} + \frac{z^2}{c^2 + v},$$

the integral

$$V = \int F(w)(a^2 + v)^{-\alpha-1}(b^2 + v)^{-\beta-1}(c^2 + v)^{-\gamma-1}dv$$

may be a solution of the partial differential equation

$$0 = V_{xx} + V_{yy} + V_{zz} + (2\alpha + 1)x^{-1}V_x + (2\beta + 1)y^{-1}V_y + (2\gamma + 1)z^{-1}V_z$$

if the contour starts from ∞ and surrounds one of the points $-a^2$, $-b^2$, $-c^2$ and returns to ∞. The restrictions are that α, β, and γ must be such that the integral is convergent and the arbitrary function F must be such that differentiations under the integral sign are permissible.

California Institute of Technology

SUR LE SYSTÈME DE SOUSLIN D'ENSEMBLES DANS L'ESPACE TRANSFINI

ISAIE MAXIMOFF

Soit Ω_i le premier nombre transfini que précèdent $\aleph_i$ nombres transfinis. La suite de nombres x_α, $(1 \leqq x_\alpha < \Omega_i)$:

$$x = \{x_1, x_2, x_3, \cdots, x_\alpha, \cdots\}, \qquad 1 \leqq \alpha < \Omega_i, \tag{1}$$

sera nommée point de l'espace transfini d'ordre i. Nous le désignons par $I_x^{(i)}$.

Appelons *espace transfini à m dimensions* l'ensemble de tous les systèmes $[x^{(1)}, x^{(2)}, \cdots, x^{(m)}]$ où $x^{(s)}$, $(1 \leqq s \leqq m)$, est un point quelconque de $I_{x^{(s)}}^{(i)}$. Cet espace sera désigné par $I_{x^{(1)}x^{(2)}\ldots x^{(m)}}^{(i)}$ ou par $I_X^{(i)}$ où $X = [x^{(1)}x^{(2)} \cdots x^{(m)}]$.

Une suite d'ensembles

$$E_1, E_2, E_3, \cdots, E_\alpha, \cdots, \qquad \alpha < \Omega_i, \tag{2}$$

de points de $I_X^{(i)}$ est appelée *convergente* (Ω_i) si pour chaque point X de cet espace il existe une valeur $\alpha = \alpha_0$, dépendant de X, telle que X appartient nécessairement à tous les E_α, $(\alpha \geqq \alpha_0)$, ou bien n'appartient à aucun des E_α, $(\alpha \geqq \alpha_0)$.

L'ensemble E de tous les points X qui appartiennent à tous les E_α, $(\alpha \geqq \alpha_0)$ (α_0 dépend de X), est appelé *limite* (Ω_i) de la suite convergente (2) et nous écrivons $E = \lim_{\alpha \to \Omega_i} E_\alpha$.

Soit

$$E_1, E_2, E_3, \cdots, E_\alpha, \cdots, \qquad \alpha < \Omega_i, \tag{3}$$

une suite quelconque d'ensembles situés dans $I_X^{(i)}$. Nous dirons que l'ensemble $S = E_1 + E_2 + E_3 + \cdots + E_\alpha + \cdots$, $(\alpha < \Omega_i)$, est *somme* (Ω_i) *et l'ensemble* $P = E_1 \cdot E_2 \cdot E_3 \cdot \cdots \cdot E_\alpha \cdot \cdots$, $(\alpha < \Omega_i)$, est *partie commune* (Ω_i) des ensembles de la suite (3).

Il est évident que l'opération $\lim_{\alpha \to \Omega_i} E_\alpha$ se ramène encore aux sommes (Ω_i) et aux parties communes (Ω_i) par la formule suivante:

$$\begin{aligned} \lim_{\alpha \to \Omega_i} E_\alpha = {} & (E_1 \cdot E_2 \cdot E_3 \cdot \cdots \cdot E_\alpha \cdot \cdots) \\ & + (E_2 \cdot E_3 \cdot E_4 \cdot \cdots \cdot E_\alpha \cdot \cdots) + \cdots \\ & + (E_\beta \cdot E_{\beta+1} \cdot E_{\beta+2} \cdot \cdots) + \cdots \end{aligned} \tag{4}$$

où $\alpha < \Omega_i$, $\beta < \Omega_i$.

Système initial. Prenons un système $\mathfrak{M}$ d'ensembles quelconques M de l'espace $I_X^{(i)}$. Nous dirons que $\mathfrak{M}$ est le système initial.

Tout ensemble E de $I_X^{(\nu)}$ qu'on obtient, à partir des ensembles du système initial au moyen des deux opérations: somme (Ω_i) et partie commune (Ω_i) répétées au plus une infinité $\aleph_i$ de fois, est dit ensemble du système B (de Borel) ou ensemble de Borel.

Soit $\nu=(n_1, n_2, n_3, \cdots, n_\kappa, \cdots)$, $\kappa<\Omega_0$, une suite quelconque de nombres n_κ tels que $1\leqq n_\kappa<\Omega_i$. Désignons par $\{\nu\}$ l'ensemble de toutes les suites ν.

Soit $F(n_1)$, $F(n_1n_2)$, $F(n_1n_2n_3)$, $\cdots$, $F(n_1n_2\cdots n_\kappa)$, $\cdots$ une suite d'ensembles quelconques de $I_X^{(\nu)}$. Nous introduisons la désignation

$$E(\nu) = F(n_1)\cdot F(n_1n_2)\cdot F(n_1n_2n_3)\cdot\cdots\cdot F(n_1n_2\cdots n_\kappa)\cdot\cdots .$$

Si A est la somme de tous les ensembles $E(\nu)$, ν parcourant tous les éléments de $\{\nu\}$, convenons écrire $A=\mathfrak{S}_\nu E(\nu)$ ou

$$(5)\quad A = \mathop{\mathfrak{S}}_{\nu} F(n_1)\cdot F(n_1n_2)\cdot F(n_1n_2n_3)\cdot\cdots\cdot F(n_1n_2\cdots n_\kappa)\cdot\cdots,\ \kappa<\Omega_0.$$

Nous dirons que l'ensemble A appartient au système de Souslin si tous les ensembles $F(n_1)$, $F(n_1n_2)$, $F(n_1n_2n_3)$, $\cdots$, $F(n_1n_2\cdots n_\kappa)$, $\cdots$ sont contenus dans le système B.

Sans restreindre la généralité de définition on peut supposer ici que

$$(6)\qquad F(n_1)\supset F(n_1n_2)\supset F(n_1n_2n_3)\supset\cdots\supset F(n_1n_2\cdots n_\kappa)\supset\cdots .$$

Il est évident que tout ensemble du système B appartient au système de Souslin.

Par le raisonnement parfaitement analogue au raisonnement classique sur le système de Souslin dans l'espace ordinaire (voir Dr. F. Hausdorff, *Mengenlehre*, Berlin, 1927, pp. 92–93) on peut démontrer le théorème suivant.

THÉORÈME 1. *Si les ensembles* $F(n_1, n_2, \cdots, n_\kappa)$, $(1\leqq\kappa<\Omega_0)$, *dans la formule* (5) *appartiennent au système de Souslin, l'ensemble* A *l'est aussi.*

En effet, soient $F(n_1n_2\cdots n_\kappa)=\mathfrak{S}_{\alpha^{(\kappa)}}H_{\alpha^{(\kappa)}}$ où

$$H_{\alpha^{(\kappa)}} = M_{\alpha_1^{(\kappa)}}^{n_1n_2\cdots n_\kappa}\cdot M_{\alpha_2^{(\kappa)}}^{n_1n_2\cdots n_\kappa}\cdot M_{\alpha_3^{(\kappa)}}^{n_1n_2\cdots n_\kappa}\cdot\cdots$$

et les ensembles $M_{\alpha_r^{(\kappa)}}^{n_1n_2\cdots n_\kappa}$ appartiennent au système de Borel. Maintenant nous rangeons les éléments de la matrice

$$A = \left\|\begin{array}{cccc} n_1 & n_2 & n_3\cdots \\ \alpha_1^1 & \alpha_2^1 & \alpha_3^1\cdots \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2\cdots \\ \cdot & \cdot & \cdot\ \cdots \end{array}\right\|$$

en suite simple: $m_1, m_2, m_3, \cdots, m_s, \cdots, (s<\Omega_0)$, de telle manière que la matrice A soit égale à la matrice

$$B=\left\|\begin{array}{cccc} m_1 & m_3 & m_5 & \cdots \\ m_2 & m_6 & m_{10} & \cdots \\ m_4 & m_{12} & m_{20} & \cdots \\ m_8 & m_{24} & m_{40} & \cdots \\ \cdot & \cdot & \cdot & \cdots \end{array}\right\|.$$

Après cela nous introduisons la désignation

$$\text{(a)}\qquad M[2^r(2\kappa+1)]=M^{m_1m_3\cdots m_{2^r+1}}_{m_{2^r+1}m_{2^r+1}{}_3m_{2^r+1}{}_5\cdots m_{2^r+1}{}_{(2\kappa+1)}},\qquad r=0,1,2,3,\cdots,\ \kappa=0,1,2,3,\cdots.$$

De cette formule il suit que à tout ensemble $M[s]$ correspond un et un seul système $(m_1, m_2, \cdots, m_{2_s})$ et inversement à tout système $(m_1, m_2, \cdots, m_{2_s})$ correspond un et un seul ensemble $M[s]$. C'est pour cette raison on peut introduire la désignation nouvelle

$$\text{(b)}\qquad M_{m_1m_2\cdots m_{2_s}}=M[s].$$

Maintenant nous introduisons l'ensemble $M_{m_1m_2\cdots m_{2s}m_{2s+1}}$ d'après

$$\text{(c)}\qquad M_{m_1m_2\cdots m_{2s+1}}=M_{m_1m_2\cdots m_{2s}}.$$

Ceci étant, on peut écrire la formule

$$A=\mathfrak{S}_\mu F(n_1)F(n_1n_2)\cdots F(n_1n_2\cdots n_\kappa)\cdots=\mathfrak{S}_\mu M_{m_1}M_{m_1m_2}M_{m_1m_2m_3}\cdots,$$

où $\mu=(m_1m_2m_3\cdots)$. Comme $M_{m_1m_2\cdots m_\kappa}$ appartient au système de Borel, l'ensemble A appartient au système de Souslin, ce qu'il fallait démontrer.

Voici les deux théorèmes qui sont conséquences immédiates du Théorème 1.

THÉORÈME 2. *Si les ensembles E_α, $(1\leqq\alpha<\Omega_i)$, appartiennent au système de Souslin, la somme (Ω_i): $E_1+E_2+E_3+\cdots+E_\alpha+\cdots$ l'est aussi.*

En effet, si nous posons $F(n_1n_2\cdots n_\kappa)=E_{n_1}$, nous obtenons:

$$\mathfrak{S}F(n_1)\cdot F(n_1n_2)\cdot F(n_1n_2n_3)\cdot\cdots\cdot F(n_1n_2n_3\cdots n_\kappa)\cdot\cdots = E_1+E_2+E_3+\cdots+E_{n_1}+\cdots,\quad n_1<\Omega_i,\ \kappa<\Omega_0.$$

THÉORÈME 3. *Si les ensembles E_n, $(1\leqq n<\Omega_0)$, appartiennent au système de Souslin, l'ensemble $E_1\cdot E_2\cdot E_3\cdot\cdots\cdot E_n\cdot\cdots$ l'est aussi.*

En effet, si nous posons $F(n_1 n_2 \cdots n_\kappa) = E_\kappa$, nous obtenons:

$$\mathop{\mathfrak{S}}_{\rho} F(n_1) \cdot F(n_1 n_2) \cdot \cdots \cdot F(n_1 n_2 \cdots n_\kappa) \cdot \cdots = E_1 \cdot E_2 \cdot E_3 \cdot \cdots \cdot E_\kappa \cdot \cdots, \qquad \kappa < \Omega_0.$$

Si un ensemble A appartient au système de Souslin, nous l'appelons ensemble de Souslin.

Soit A un ensemble de Souslin quelconque. Dans ce cas il y a lieu la formule (5). Nous introduisons la désignation $r = (n_1, n_2, \cdots, n_\kappa)$, $(\kappa < \Omega_0)$, et nous dirons que le système r *précède* le système $(r, n) = (n_1, n_2, \cdots, n_\kappa, n)$ ou le système (r, n) *suit* le système (r).

Soit R_0 un ensemble quelconque de systèmes $(n_1, n_2, \cdots, n_\kappa)$ où $1 \leqq n_\tau < \Omega_i$; $\tau = 1, 2, 3, \cdots, \kappa$; $\kappa < \Omega_0$. Si $r = (n_1 n_2 \cdots n_\kappa)$ appartient à R_0 et s'il n'y a aucun élément (r, n) qui suit l'élément r et appartient à R_0, nous dirons que r est un élément *final* de R_0.

Étant donné un ensemble quelconque R_0 d'éléments r, nous formons la suite des ensembles R_α provenant de l'ensemble donné R_0:

$$R_0, R_1, R_2, \cdots, R_\alpha, \cdots, R_\gamma, \cdots \tag{7}$$

de manière suivante. On obtient R_γ, si γ est de première espèce, $\gamma = \gamma^* + 1$, en enlevant de l'ensemble précédent déjà défini R_{γ^*} tout élément final, et, si γ est de seconde espèce, on obtient R_γ en faisant la partie commune des ensembles précédents $R_{\gamma'}$, $(\gamma' < \gamma)$, déjà définis.

D'après ce qui précède l'inégalité $\gamma < \gamma'$ entraîne $R_\gamma \supset R_{\gamma'}$. Désignons par Δ_γ l'ensemble des éléments de $R_{\gamma+1}$ sans faire partie de R_γ. On voit que la condition nécessaire et suffisante pour que Δ_γ soit nul est que $R_{\gamma+1} = R_\gamma$. Chacun des ensembles

$$\Delta_0, \Delta_1, \Delta_2, \Delta_3, \cdots, \Delta_\gamma, \cdots \tag{8}$$

à la puissance égale au plus à $\aleph_i$, et deux d'entre eux n'ont aucun élément commun. Il y a donc au plus une infinité $\aleph_i$ d'ensembles (8) non nuls; les indices γ de ceux des ensembles (8) qui ne sont pas nuls forment ainsi un ensemble de la puissance égale au plus à $\aleph_i$, par suite, il y a un nombre ρ, $(\rho < \Omega_{i+1})$, qui a la propriété suivante: si $\rho' \geqq \rho$, on a $\Delta_{\rho'} = 0$ d'où résulte

$$R_\rho = R_{\rho+1} = R_{\rho+2} = \cdots = R_{\rho'} = \cdots, \qquad \rho' > \rho. \tag{9}$$

Désignons par $R_{\Omega_{i+1}}$ la partie commune à tous les ensembles (7). On voit bien que les ensembles (7) sont, pour les valeurs suffisamment grandes de l'indice, identiques à $R_{\Omega_{i+1}}$.

Il y a un nombre bien déterminé qui est le plus petit de tels qu'on

ait $R_\eta = R_{\Omega_{i+1}}$. Nous le désignons par η et dirons que η est l'indice de R_0: pour $\xi < \eta$ on a $R_\xi \supset R_{\xi+1}$ et pour $\xi \geqq \eta$ on a $R_\xi = R_{\Omega_i}$.

Désignons par $K(R_0)$ l'ensemble $R = R_\eta = R_{\eta+1} = R_{\eta+2} = \cdots$ et par $R_0(X)$ l'ensemble de tous les systèmes $r = (n_1 n_2 \cdots n_\kappa)$, $\kappa < \Omega_0$, tels que X appartient à l'ensemble $F(r) = F(n_1 n_2 \cdots n_\kappa)$. Soit η le plus petit nombre tel que $KR_0(X) = R_\eta(X) = R_{\eta+1}(X) = \cdots$. Dans ce cas η est dit *l'indice* du point X, et nous écrivons l'égalité $\eta = \eta(X)$.

D'ailleurs, $R(X) = R_\eta(X)$ est aussi appelé indice du point X. Désignons par A_η (B_η) l'ensemble de tous les points X appartenant à l'ensemble A (B) où $B = CA$ et ayant l'indice η. Alors nous avons les formules

$$A = A_0 + A_1 + A_2 + \cdots + A_\gamma + \cdots,$$
$$B = B_0 + B_1 + B_2 + \cdots + B_\gamma + \cdots, \quad 0 \leqq \gamma < \Omega_{i+1}.$$

Il est évident que les ensembles

$$A_0, A_1, A_2, A_3, \cdots, A_\gamma, \cdots, \qquad B_0, B_1, B_2, B_3, \cdots, B_\gamma, \cdots, \quad 0 \leqq \gamma < \Omega_{i+1},$$

n'ont aucun point commun deux à deux. Nous les appelons *constituantes* respectivement de A et de B.

Maintenant nous allons démontrer notre théorème fondamental suivant.

THÉORÈME 4. *Si A est un ensemble de Souslin quelconque, alors nous avons les développements*

$$A = A_0 + A_1 + A_2 + \cdots + A_\alpha + \cdots,$$
$$B = B_0 + B_1 + B_2 + \cdots + B_\alpha + \cdots, \qquad \alpha < \Omega_{i+1},$$

de l'ensemble A et de son complémentaire $B = CA$ en une infinité $\aleph_{i+1}$ d'ensembles constituantes A_α, B_α, $(0 \leqq \alpha < \Omega_{i+1})$, chacun desquels appartient au système B.

Démonstration de ce théorème consiste de quatre parties.

La première partie. Tout d'abord montrons que l'ensemble $R_\xi(X)$ défini précédemment contient tous les éléments précédents, c'est à dire,

$$(10) \qquad (r, n) \in R_\xi(X) \to (r) \in R_\xi(X),$$

où $\to$ signifie l'implication logique.

Cette proposition est vraie pour $\xi = 0$.

En effet, si (r, n) appartient à $R_0(X)$, on a $X \in F(r, n)$, par suite, $X \in F(r)$ d'où il suit $r \in R_0(X)$. Maintenant supposons que cette prop-

osition est vraie pour $\xi=\alpha$ et démontrons la pour $\xi=\alpha+1$; d'après l'hypothèse faite $r \in R_\alpha(X)$. Donc, $R_\alpha(X)$ contient (r) et (r, n), par suite, (r) n'est pas un élément final de $R_\alpha(X)$; il en suit que $(r) \in R_{\alpha+1}(X)$.

Soit η un nombre quelconque de seconde espèce, $\eta<\Omega_{i+1}$, et soit $(r, n) \in R_\eta(X)$; alors (r, n) est contenu dans chaque $R_\xi(X)$ où $\xi<\eta$.

Si $(r, n) \in R_\xi(X)$, on a $(r) \in R_\xi(X)$ pour tout $\xi<\eta$, par suite, $r \in R_\eta(X)$.

La deuxième partie. Désignons par $F_\xi(r)$ l'ensemble des points X tels que $r \in R_\xi(X)$. Cela veut dire que la relation $r \in R_\xi(X)$ est équivalente à la relation $X \in F_\xi(r)$. Si $X \in F_\xi(r, n)$ on a $(r, n) \in R_\xi(X)$, par suite, $r \in R_\xi(X)$ d'où il suit $X \in F_\xi(r)$. Donc,

$$F_\xi(r, n) \subset F_\xi(r). \tag{11}$$

Il est évident que

$$F_0(r) = F(r), \quad F_{\xi+1}(r) = \mathfrak{S}_n F_\xi(r, n), \quad F_\eta(r) = \mathfrak{D}_{\xi<\eta} F_\xi(r), \quad n < \Omega_i, \tag{12}$$

où $\mathfrak{S}_n$ est le signe de la sommation s'étendant sur tous $n<\Omega_i$ et $\mathfrak{D}_{\xi<\eta}$ est le signe de la partie commune s'étendant sur tous $\xi<\eta$.

La première des égalités (12) est une consequence immédiate de la définition de $R_0(X)$. En effet, la relation $X=F_0(r)$ est équivalente à la relation $r \in R_0(X)$. Mais cette dernière est équivalente à la relation $X \in F(r)$. La deuxième des égalités (12) s'obtient par le raisonnement suivant: la relation $X \in F_{\xi+1}(r)$ est équivalente à la relation $r \in R_{\xi+1}(X)$. Mais dans ce cas r n'est pas élément final de $R_\xi(X)$, par suite, $(r, n) \in R_\xi(X)$ d'où $X \in F_\xi(r, n)$.

La troisième des égalités s'obtient par le raisonnement analogue: la relation $X \in F_\eta(r)$ est équivalente à la relation $r \in R_\eta(X) = \mathfrak{D}_{\xi<\eta} R_\xi(X)$. Mais la dernière est équivalente à la suite des relations:

$$r \in R_0(X), r \in R_1(X), \cdots, r \in R_\xi(X), \cdots, \qquad \xi < \eta,$$

équivalentes respectivement aux relations

$$X \in F_0(r), X \in F_1(r), \cdots, X \in F_\xi(r), \cdots, \qquad \xi < \eta.$$

Mais le dernier système des relations est équivalent à la relation $X \in \mathfrak{D}_{\xi<\eta} F_\xi(r)$ d'où il suit $F_\eta(r) = \mathfrak{D}_{\xi<\eta} F_\xi(r)$.

La troisième partie. Maintenant posons

$$S_\xi = \mathfrak{S}_r F_\xi(r), \tag{13}$$

$$T_\xi = \mathfrak{S}_r [F_\xi(r) - F_{\xi+1}(r)]. \tag{14}$$

$F_0(r)$ est ensemble de Borel, puisque $F_0(r)=F(r)$. En partant des

ensembles $F_0(r)$ et en tenant compte de formules (12) on peut démontrer que tous les ensembles $F_\xi(r)$ appartiennent au système B, par suite, S_ξ et T_ξ appartiennent aussi à ce système.

Il est à remarquer que S_ξ est l'ensemble des points X pour lesquels $R_\xi(X) \neq 0$, et T_ξ est l'ensemble des points X pour lesquels

$$R_\xi(X) - R_{\xi+1}(X) \neq 0.$$

La quatrième partie. Maintenant nous allons démontrer que

$$S_\xi = A + \sum_{\eta>\xi} B_\eta, \qquad T_\xi = \sum_{\eta>\xi} A_\eta + \sum_{\eta>\xi} B_\eta. \tag{15}$$

Pour cela nous faisons les remarques suivantes: En premier lieu, si $X \in A$ on a $KR_0(X) \neq 0$ et inversement si $KR_0(X) \neq 0$ on a $X \in A$. En second lieu, si $X \in B$ on a $KR_0(X) = 0$ et inversement si $KR_0(X) = 0$ on a $X \in B$. En troisième lieu, $R_\xi(X) \supset R_{\xi+1}(X)$ pour $\xi < \eta(X)$ et $R_\xi(X) = R_{\xi+1}(X)$ pour $\xi \geqq \eta(X)$ où $\eta(X)$ est l'indice du point X.

Soit X un point quelconque de $I_X^{(i)}$. Nous considérons les cas possibles suivants:

Le premier cas: $X \in S_\xi$. Dans ce cas $R_\xi(X) \neq 0$. Inversement, si $R_\xi(X) \neq 0$ on a $X \in S_\xi$. Si $X \in A$ on a $KR_0(X) \neq 0$, par suite, $R_\xi(X) = 0$ d'où $X \in S_\xi$, c'est à dire, $A \subset S_\xi$. Si $X \in B$ on a $KR_0(X) = 0$, par suite, $R_\eta(X) = 0$, mais $X \in S_\xi$, par suite, $X \in R_\xi(r)$ d'où $\xi < \eta$, c'est à dire, l'indice de X est supérieure à ξ. Donc $S_\xi = A + \sum_{\eta>\xi} B_\eta$.

Le second cas: $X \in T_\xi$. Dans ce cas $R_\xi(X) - R_{\xi+1}(X) \neq 0$, alors $\eta > \xi$, par suite, $T_\xi = \sum_{\eta>\xi} A_\eta + \sum_{\eta>\xi} B_\eta$. De formules (15) on tire

$$\begin{aligned} S_\xi - T_\xi &= A_0 + A_1 + A_2 + \cdots + A_\xi, \\ I_x^{(i)} - S_\xi &= B_0 + B_1 + B_2 + \cdots + \beta_\xi. \end{aligned} \tag{16}$$

Au moyen de ces formules par le procédé d'induction on peut démontrer que les ensembles A_ξ et B_ξ sont contenus dans le système B, ce qu'il fallait démontrer.

Maintenant posons les définitions suivantes:

(a) Si pour un point $x = \{x_1, x_2, x_3, \cdots, x_\alpha, \cdots\}$, $(\alpha < \Omega_i)$, de l'espace $I_x^{(i)}$ tous les nombres x_n, $(1 \leqq n < \Omega_0)$, sont finis, $x_n < \Omega_0$, alors nous dirons que x possède un *noyau* $\bar{x} = \{x_1, x_2, x_3, \cdots, x_n, \cdots\}$, $(1 \leqq n < \Omega_0)$, qui est un point de l'espace de Baire $I^{(0)}$.

(b) Désignons par $n(E)$ l'ensemble formé des noyaux de tout les points de E; $n(E)$ est dit le noyau de l'ensemble E.

(c) Si E_0 est le noyau d'un ensemble E de Borel, nous dirons que E_0 est ensemble de Borel $(\aleph_i)$ dans $I_x^{(0)}$.

(d) Si E_0 est le noyau d'un ensemble E de Souslin, nous dirons que E_0 est ensemble de Souslin $(\aleph_i)$ dans $I_x^{(0)}$.

(e) Si E contient tout ensemble $\mathcal{E}$ tel que $n(E)=n(\mathcal{E})$, nous dirons que E est *normal*.

Ceci étant, prenons un ensemble de Souslin quelconque A de points de $I_x^{(0)}$. D'après le Théorème 4 on a

$$A = A_0 + A_1 + A_2 + \cdots + A_\alpha + \cdots,$$
$$CA = B_0 + B_1 + B_2 + \cdots + B_\alpha + \cdots, \quad 0 \leqq \alpha < \Omega_{i+1}.$$

Il est évident que si E est normal on a

$$n(CE) = C[n(E)] \tag{17}$$

d'où il suit pour A normal

$$\begin{aligned} n(A) &= n(A_0) + n(A_1) + n(A_2) + \cdots + n(A_\alpha) + \cdots, \\ C[n(A)] &= n(B_0) + n(B_1) + n(B_2) + \cdots + n(B_\alpha) + \cdots, \\ &\qquad\qquad\qquad\qquad\qquad\qquad \alpha < \Omega_{i+1}. \end{aligned} \tag{18}$$

Donc, nous avons obtenu le théorème suivant.

THÉORÈME 5. *Pour tout ensemble a de Souslin* $(\aleph_i)$ *dans* $I_x^{(0)}$ *on peut trouver les ensembles de Borel* $(\aleph_i)$:

$$a_0, a_1, a_2, \cdots, a_\alpha, \cdots, \qquad b_0, b_1, b_2, \cdots, b_\alpha, \cdots, \qquad \alpha < \Omega_{i+1},$$

de cet espace tels que

$$a = a_0 + a_1 + a_2 + \cdots + a_\alpha + \cdots,$$
$$b = b_0 + b_1 + b_2 + \cdots + b_\alpha + \cdots, \qquad \alpha < \Omega_{i+1}.$$

TCHEBOKSARY, U.R.S.S.

ON LINEAR POLYGON TRANSFORMATIONS[1]

JESSE DOUGLAS

1. **Introduction, definitions.** In a recent paper[2] the author has developed a general theory of linear transformations of polygons, regarded as lying in the complex plane. By a *polygon* we understand a system of points, or complex numbers: $(z_1, z_2, \cdots, z_n)$, called *vertices*, which are taken in a definite *cyclic* order. This is to say that two n-gons, $(z_1, z_2, \cdots, z_n)$ and $(w_1, w_2, \cdots, w_n)$, are the same when and only when $z_i = w_{i+k}$ for $i = 1, 2, \cdots, n$, where k has any fixed one of the values $0, 1, 2, \cdots, n-1$ and all indices are taken modulo n. To be distinguished from a polygon is a *multipoint*, where the definition of equivalence is the identity of corresponding points in the order given: $z_i = w_i$ for $i = 1, 2, \cdots, n$.

A linear multipoint transformation is simply the general linear transformation of n complex variables:

$$z_i' = \sum_{j=1}^{n} a_{ij} z_j, \qquad i = 1, 2, \cdots, n, \tag{1}$$

where the coefficients a_{ij} may be any complex numbers. For a linear polygon transformation, on the other hand, a certain cyclicity is required: a cyclic permutation of the z's must produce the same cyclic permutation of the z''s. Thus we may write down arbitrarily the first line of an L.P.T.:[3]

$$z_1' = \alpha_0 z_1 + \alpha_1 z_2 + \cdots + \alpha_{r-1} z_r,$$

---

[1] Presented to the Society, October 29, 1938, under the title *Geometry of polygons in the complex plane.*

[2] *Geometry of polygons in the complex plane,* Journal of Mathematics and Physics, vol. 19 (1940), pp. 93–130, and this Bulletin, abstract 44-9-390. See, as a preliminary to the present theory, papers by E. Kasner and his students in Scripta Mathematica, vol. 2 (1934), pp. 131–138, and vol. 4 (1936), pp. 37–49. Kasner considers the polygon derived from a given one by taking the midpoint of each side, and, more generally, by taking the centroid of r consecutive vertices. These are special linear polygon transformations (2), where $\alpha_0 = \alpha_1 = \cdots = \alpha_{r-1} = 1/r$. Kasner uses real cartesian coordinates, and his polygons may lie in euclidean space of any number of dimensions.

The basic ideas of the present paper are (i) to regard the polygons as lying in the *complex plane,* (ii) to consider general linear polygon transformations (2) with any *complex* coefficients. As pointed out in §1, this is equivalent to taking a "center of gravity" with complex "weights."

The author delivered a series of lectures on the present topic at Columbia University in July, 1939.

[3] L.P.T. denotes "linear polygon transformation" throughout this paper.

but then the whole transformation is determined by cyclic permutation, that is:

$$(2) \qquad z_k' = \alpha_0 z_k + \alpha_1 z_{k+1} + \cdots + \alpha_{r-1} z_{k+r-1}, \qquad k = 1, 2, \cdots, n,$$

where the indices of z are to be taken modulo n. Here r may have any value from 1 to n inclusive, and we may suppose $\alpha_0 \neq 0$,[4] $\alpha_{r-1} \neq 0$, the transformation being then called *r-ary*. If $r < n$, we may fill out formula (2) to the general form (1) by using zero coefficients in each line of (2) for those z's which do not effectively appear.

Formula (2) represents the general *linear polygon transformation L.*[5] Its characteristic property is to be transformed into itself by a cyclic permutation $C = (z_1, z_2, \cdots, z_n)$ of the z's: $C^{-1}LC = L$, or to be commutative with such a permutation: $CL = LC$. Since $C^{-1}LC = L$ and $C^{-1}L'C = L'$ imply by multiplication $C^{-1}LL'C = LL'$, it follows that *the product of two L.P.T.'s is again an L.P.T.* Thus, all L.P.T.'s for a given value of n form a group,[6] a subgroup G_n of n complex parameters in the total linear group G_{n^2} of n^2 complex parameters represented by (1). In §2 another proof of the same fact will appear (Theorem II).

Particularly interesting are the special L.P.T.'s, to be denoted by M, for which

$$(3) \qquad \alpha_0 + \alpha_1 + \cdots + \alpha_{r-1} = 1.$$

As is immediately verifiable, (3) is the necessary and sufficient condition that a given L.P.T. (2) be permutable with an arbitrary similitude transformation $S: z' = Az + B$ (A, B complex); that is, $MS = SM$, or $S^{-1}MS = M$, for every S and every M. It follows, as in the preceding paragraph, that *the transformations M form a group*, a subgroup G_{n-1} of the group G_n of all L.P.T.'s.

If the polygon P' is the image of the polygon P by any transformation M, the remarks just made state that the relation between P and P' is invariant under an arbitrary similitude transformation; we may say that P' is a "similitude concomitant" of P. Accordingly, we shall term any L.P.T., M, which obeys the condition (3) a *similitude*

[4] Evidently we can always bring about $\alpha_0 \neq 0$ by cyclic renumbering of the vertices z.

[5] We dispense with the consideration of nonhomogeneous L.P.T.'s, since these have the same constant term in each line and therefore differ from a homogeneous L.P.T. only by a translation.

[6] It should be emphasized that here, and throughout this paper, we use the term "group" only in the sense of possession of the specific group property: closure with respect to composition of elements. Since the determinant of an L.P.T. may be zero, an inverse transformation does not always exist.

construction on the polygon $P=(z_1, z_2, \cdots, z_n)$ and a *binary, ternary, $\cdots$, r-ary* S.C.[7] according as two, three, $\cdots$, r consecutive vertices of P are effectively involved in the formula (2) of the transformation ("effectively" meaning $\alpha_0 \neq 0, \alpha_{r-1} \neq 0$).

If furthermore the α's are *real*, then, as is easily verified, we have for every affine transformation a, that is, $z'=Az+B\bar{z}+C$ (A, B, C complex, $\bar{z}$ the conjugate of z), the relation $a^{-1}Ma=M$ or $Ma=aM$. The polygon P' is then an "affine concomitant" of the polygon P; accordingly, for *real* values of the α's obeying (3), we term M an *affine construction*.

Evidently, under the condition (3), we can consider that (2) defines z_k' as a kind of "center of gravity" of $z_k, z_{k+1}, \cdots, z_{k+r-1}$ with *complex* "weights" proportional to $\alpha_0, \alpha_1, \cdots, \alpha_{r-1}$. In the case of an affine construction, where the weights α are real, we have an actual center of gravity in the ordinary sense.

We may observe here, as in our cited paper,[8] that the centroid of any polygon is invariant under any S.C., M. For by addition of all the equations (2),

$$z_1' + z_2' + \cdots + z_n' = (\alpha_0 + \alpha_1 + \cdots + \alpha_{r-1})(z_1 + z_2 + \cdots + z_n); \tag{3a}$$

then take account of (3) and divide by n.

Of special importance is the *binary* S.C., M_2:

$$z_k' = \alpha_0 z_k + \alpha_1 z_{k+1}, \qquad k = 1, 2, \cdots, n, \tag{4}$$

$$\alpha_0 + \alpha_1 = 1. \tag{4'}$$

The point z_k' is here determined by the simple geometric condition that *the triangle* (z_k, z_{k+1}, z_k') *is directly*[9] *similar to the fixed triangle* $(-\alpha_1, \alpha_0, 0)$. This is seen by considering the equations in A, B

$$z_k = -\alpha_1 A + B, \quad z_{k+1} = \alpha_0 A + B, \quad z_k' = B, \tag{5}$$

whose solvability is the condition for the existence of a similitude transformation $z'=Az+B$ converting the triangle $(-\alpha_1, \alpha_0, 0)$ into (z_k, z_{k+1}, z_k'). By use of (4'), the solution of the first two of these equations is

$$A = z_{k+1} - z_k, \; B = \alpha_0 z_k + \alpha_1 z_{k+1};$$

[7] S.C. is "similitude construction" throughout this paper.

[8] In this way we shall always refer to the paper whose title is given in the first footnote.

[9] That is, with preservation of sense.

and then equation (4) expresses exactly that the third equation (5) is satisfied.

If α_0, α_1 are real, the point z_k' lies upon the side z_kz_{k+1} and divides it in the fixed ratio $z_kz_k' : z_k'z_{k+1} = \alpha_1 : \alpha_0$.

2. L.P.T.'s **and cyclic matrices.** The matrix of the general L.P.T. is a *cyclic matrix*

$$L = \begin{Vmatrix} \alpha_0 & \alpha_1 & \cdot & \cdots & \alpha_{n-2} & \alpha_{n-1} \\ \alpha_{n-1} & \alpha_0 & \alpha_1 & \cdots & \cdot & \alpha_{n-2} \\ \alpha_{n-2} & \alpha_{n-1} & \alpha_0 & \cdots & \cdot & \alpha_{n-3} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \alpha_1 & \alpha_2 & \cdot & \cdots & \alpha_{n-1} & \alpha_0 \end{Vmatrix}, \tag{6}$$

where the circular derivation of each row from the preceding is the characteristic feature. Any number of the α's may be zero, so that the L.P.T. can be r-ary with $r = 1, 2, \cdots, n$. We shall speak interchangeably of an L.P.T., L, and its corresponding cyclic matrix L.

The simplest L.P.T. is a cyclic permutation of vertices, C:

$$z_1' = z_2, \quad z_2' = z_3, \quad \cdots, \quad z_{n-1}' = z_n, \quad z_n' = z_1.$$

As a matrix, this is

$$C = \begin{Vmatrix} 0 & 1 & 0 \cdots 0 \\ 0 & 0 & 1 \cdots 0 \\ \cdot & \cdot & \cdots\cdots \\ 0 & 0 & \cdots\cdot 1 \\ 1 & 0 & \cdots\cdot 0 \end{Vmatrix}.$$

We may say that C has 1's in its second "cyclic diagonal" and 0's everywhere else. In general, as is seen at once, C^k has 1's in its $(k+1)$th cyclic diagonal and 0's elsewhere. Obviously, $C^n = I$, the identity matrix, which has 1's in its first (cyclic) diagonal and 0's elsewhere.

From these remarks we have directly by comparison with (6):

$$L = \alpha_0 I + \alpha_1 C + \alpha_2 C^2 + \cdots + \alpha_{n-1} C^{n-1}. \tag{7}$$

Conversely, every polynomial in C of this form represents a cyclic matrix, namely the one whose first, or generating, row consists of the coefficients of the polynomial in order. Hence we have the following theorem.

THEOREM I. *The necessary and sufficient condition that an n-rowed square matrix L be cyclic is that it be expressible as a polynomial of degree $n-1$ at most in the matrix C of a cyclic permutation. This mode of representation of L is unique.*

If in (7) we replace C by a (complex) numerical variable x, and I by 1, we obtain a polynomial $\phi(x)=\alpha_0+\alpha_1 x+\cdots+\alpha_{n-1}x^{n-1}$, that we shall term the *auxiliary polynomial*[10] of L.

If any two polynomials of the form (7) are multiplied together, the product can be reduced to the degree $n-1$ at most by means of the relation $C^n=I$. From this remark and Theorem I the following three theorems result immediately.

THEOREM II. *The product of two cyclic matrices is again a cyclic matrix.*

THEOREM III. *The multiplication of cyclic matrices is isomorphic with the multiplication of their auxiliary polynomials, modulo x^n-1.*

THEOREM IV. *The multiplication of cyclic matrices is commutative.*

Thus cyclic matrices form a commutative subalgebra of order n in the non-commutative algebra of order n^2 formed by general n-rowed square matrices.

The following alternative form of proof of these theorems, directly in terms of the explicit representation (6) of a cyclic matrix, may be of interest. Let α_m, $(m=0, 1, \cdots, n-1)$, denote any n complex numbers. Then a cyclic matrix of order n, $\|a_{ij}\|$, is one where $a_{ij}=\alpha_{j-i}$, *the index $j-i$ being taken modulo n*; that is, in case $j-i$ is negative, $j-i+n$ is to be used instead. If $\|b_{ij}\|$ is another cyclic matrix: $b_{ij}=\beta_{j-i}$, $j-i$ taken modulo n, then $\|a_{ij}\|\times\|b_{ij}\|=\|c_{ij}\|$, where

$$c_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj}=\sum_{k=1}^{n}\alpha_{k-i}\beta_{j-k}=\sum\alpha_s\beta_t$$

(in the last summation s, t evidently take once and only once every combination of values such that $s+t\equiv j-i$, modulo n). In other words, defining

$$\gamma_m=\sum\alpha_s\beta_t, \qquad s+t\equiv m \pmod n, \tag{8}$$

we have $c_{ij}=\gamma_{j-i}$, $j-i$ taken modulo n.

This proves Theorem II. Theorem III follows because (8) repre-

[10] This is only of degree $r-1$ if L is r-ary (see (9)). In our cited paper we defined the auxiliary polynomial as $\alpha_0x^{r-1}+\alpha_1x^{r-2}+\cdots+\alpha_{r-1}$. With this difference of notation in mind, no ambiguity should arise in comparing the two papers.

sents the law of multiplication of the auxiliary polynomials modulo x^n-1:

$$\sum_{m=0}^{n-1}\gamma_m x^m \equiv \sum_{s=0}^{n-1}\alpha_s x^s \cdot \sum_{t=0}^{n-1}\beta_t x^t \pmod{(x^n-1)}.$$

Theorem IV expresses the symmetry of (8) in the indices s, t of α, β.

As has been pointed out, all these theorems can be stated in terms of L.P.T.'s[11] instead of cyclic matrices. The auxiliary polynomial of an r-ary L.P.T. is of degree $r-1$:

$$\phi(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_{r-1}x^{r-1}. \tag{9}$$

We term $\phi(x)=0$ the *auxiliary equation* of the L.P.T.

By factorization of its auxiliary polynomial, or solution of its auxiliary equation, *every r-ary L.P.T. can be factored into* $r-1$ *binary L.P.T.'s*. This results from Theorem III; in fact, if

$$\phi(x) = (\alpha_0^{(1)} + \alpha_1^{(1)}x)(\alpha_0^{(2)} + \alpha_1^{(2)}x)\cdots(\alpha_0^{(r-1)} + \alpha_1^{(r-1)}x), \tag{10}$$

then L is equal to the product of the $r-1$ binary L.P.T.'s

$$L_2^{(j)} = \alpha_0^{(j)}I + \alpha_1^{(j)}C, \qquad j = 1, 2, \cdots, r-1.$$

The binary components $L_2^{(j)}$ are evidently not unique; rather, as is evident, each is subject to multiplication by a (complex) scalar $\lambda^{(j)}$ provided that the product of all the $\lambda^{(j)}$'s is unity.

If L is an S.C., M, so that (3) is verified, we have $\phi(1)=1$, or by (10),

$$1 = (\alpha_0^{(1)} + \alpha_1^{(1)})(\alpha_0^{(2)} + \alpha_1^{(2)})\cdots(\alpha_0^{(r-1)} + \alpha_1^{(r-1)}). \tag{11}$$

Dividing (10) by (11), we obtain

$$\phi(x) = (\beta_0^{(1)} + \beta_1^{(1)}x)(\beta_0^{(2)} + \beta_1^{(2)}x)\cdots(\beta_0^{(r-1)} + \beta_1^{(r-1)}x)$$

where $\beta_0^{(j)}=\alpha_0^{(j)}/(\alpha_0^{(j)}+\alpha_1^{(j)})$, $\beta_1^{(j)}=\alpha_1^{(j)}/(\alpha_0^{(j)}+\alpha_1^{(j)})$; hence

$$\beta_0^{(j)} + \beta_1^{(j)} = 1. \tag{12}$$

Accordingly $M_2^{(j)}=\beta_0^{(j)}I+\beta_1^{(j)}C$ is a binary S.C. for $j=1, 2, \cdots, r-1$, and we have the following theorem.

---

[11] The commutativity of L.P.T.'s seems the more noteworthy because it is certainly not evident geometrically that two binary S.C.'s are commutative. See the geometric interpretation of a binary S.C. at the end of §1.

THEOREM V. *Every r-ary S.C., M_r, is representable uniquely as the product of $r-1$ binary S.C.'s: $M_r=\prod_{j=1}^{r-1}M_2^{(j)}$.*

The uniqueness is a consequence of the fact that the components $M_2^{(j)}$ are not subject to multiplication by a scalar, because that would disturb condition (12).

An interesting illustration of Theorem V is furnished by the ternary centroidal construction

$$z_k' = \tfrac{1}{3}z_k + \tfrac{1}{3}z_{k+1} + \tfrac{1}{3}z_{k+2}.$$

Here the auxiliary polynomial factors as follows:

$$\frac{1}{3}(1 + x + x^2) = \frac{\omega - x}{\omega - 1}\cdot\frac{\omega^2 - x}{\omega^2 - 1},$$

where $\omega=e^{2\pi i/3}$ is a complex cube root of unity. By the remarks associated with (4), the component binary S.C.'s are those which employ triangles similar respectively to (1, ω, 0), (1, ω^2, 0). Hence the centroid of any three points z_1, z_2, z_3 will be arrived at as follows. On z_1z_2, z_2z_3 construct 120° isosceles triangles to the left; let z_1', z_2' be their respective vertices. On $z_1'z_2'$ construct a 120° isosceles triangle to the right; its vertex z_1'' is the centroid of z_1, z_2, z_3.

We conclude this section by observing the following factorization of a cyclic determinant:

$$\det L = \phi(1)\cdot\phi(\omega)\cdot\phi(\omega^2)\cdot\cdots\cdot\phi(\omega^{n-1}), \tag{13}$$

where $\omega=e^{2\pi i/n}$ is a primitive nth root of unity. The proof is easily given as follows. Det L is a homogeneous polynomial of degree n in the α's. By addition of all the columns of L (refer to (6)), $\phi(1)$ is seen to be a factor of det L. By using 1, ω, ω^2, $\cdots$, ω^{n-1} as multipliers of the successive columns and adding, $\phi(\omega)$ is seen to be a factor of det L. Similarly $\phi(\omega^2)$, $\cdots$, $\phi(\omega^{n-1})$ are factors. Each of these n factors is a linear polynomial in the α's. Hence (13) holds to within a numerical factor, which is verified to be 1 by comparison of the terms in α_0^n.

3. **Geometrical applications.** In our cited paper, we consider the relations

$$R_p \equiv \sum_{k=1}^{n}\omega^{(k-1)p}z_k = 0, \quad p = 1, 2, \cdots, n-1, \tag{14}$$

as applied to any polygon. Each relation $R_p=0$ expresses an intrinsic geometric property of the polygon, being invariant under any simili-

tude transformation, $z_k' = Az_k + B$; this is easily verified if account is taken of the well known equation[12] $\sum_{k=0}^{n-1} \omega^{kp} = 0$, $(p = 1, 2, \cdots, n-1)$.

A polygon is called *regular of ω^q-type*[13] if it is similar to the polygon $(1, \omega^q, \omega^{2q}, \cdots, \omega^{(n-1)q})$. A regular polygon of ω-type is simply a (convex) regular polygon in the ordinary sense. A regular polygon of ω^q-type may be star-shaped (for example, $n=5$, $q=2$: a star-shaped regular pentagon), or it may resolve into an s-gon described t times, where $st=n$ (for example, $n=6$, $q=2$: an equilateral triangle described twice, considered as a form of regular hexagon).

It was shown (loc. cit., §5) that *the criterion for a polygon to be regular of ω^q-type is that it obey the following $n-2$ conditions*:

$$R_p = 0 \qquad \text{for } p = 1, 2, \cdots, n-1, \text{ except } n-q. \tag{15}$$

Each condition $R_p=0$ may therefore be regarded as expressing a "degree of regularity," and any number $k<n-2$ of these conditions as expressing "partial regularity" of degree k.

By multiplying equation (2) by $\omega^{(k-1)p}$ and summing for $k=1, 2, \cdots, n$, we find

$$R_p' = \phi(\omega^{-p})R_p = \phi(\omega^{n-p})R_p, \tag{16}$$

or, interchanging p and $n-p$,

$$R_{n-p}' = \phi(\omega^p)R_{n-p}. \tag{16'}$$

We infer from (16) or (16′) the following two facts:

(i) *If a polygon obeys the relation $R_p=0$, this relation persists after any L.P.T.* (That is, R_p is a *relative invariant* of the group G_n of all L.P.T.'s.)

(ii) *If ω^p is a root of the auxiliary equation $\phi(x)=0$ of an L.P.T., then this L.P.T. converts every polygon P into a polygon P' obeying the relation $R_{n-p}=0$.*

A binary S.C., $M_2 = \alpha_0 I + \alpha_1 C$, $(\alpha_0 + \alpha_1 = 1)$, has the auxiliary linear polynomial $\phi(x) = \alpha_0 + \alpha_1 x$, and hence obeys the hypothesis of (ii): $\phi(\omega^p) = 0$, when and only when $-\alpha_1 : \alpha_0 = 1 : \omega^p$. By the statement following (4′), this means that the triangle (z_k, z_{k+1}, z_k') is similar to $(1, \omega^p, 0)$. Then z_k' is at the vertex of an isosceles triangle with vertex angle $2p\pi/n$ based on $z_k z_{k+1}$ and to the left[14] of this base. We shall denote this construction whereby the polygon (z_k') is derived from the polygon (z_k) by $\Delta(2p\pi/n)$.

[12] ω^p is a root, not 1, of $x^n - 1 \equiv (x-1)(1 + x + x^2 + \cdots + x^{n-1}) = 0$.

[13] As defined, loc. cit., §2, Definition (iii).

[14] With this interpretation: an isosceles triangle to the left with vertex angle $\pi + \theta$ is an isosceles triangle to the right with vertex angle $\pi - \theta$.

We may now put (i), (ii), and the criterion (15) together so as to give the next theorem:

THEOREM VI. *If, on an* arbitrary *polygon P_0, the $n-2$ constructions $\Delta(2p\pi/n)$ for $p=1, 2, \cdots, n-1$ except q, are performed successively in any order, giving the series of polygons $P_1, P_2, \cdots, P_{n-2}$, then P_{n-2} is* regular *of ω^q-type. Further, P_{n-2} is independent of the order of these constructions. Also, the centroid of P_{n-2} and of all the intermediate polygons $P_1, P_2, \cdots$ is the same as that of P_0.*[15]

(The last statement refers to the remark associated with (3a).)

PROOF. Each operation $\Delta(2p\pi/n)$ confers the property $R_{n-p}=0$ on the new polygon, according to (ii); and the subsequent operations allow the polygon to keep this property, according to (i); therefore the final polygon P_{n-2} has all the properties (15) requisite for regularity of ω^q-type.

The case $n=3$ of Theorem VI is the following well known theorem of elementary geometry: If on each side of an *arbitrary* triangle as base, a 120° isosceles triangle is constructed, always outward or always inward, then the vertices of these isosceles triangles form an *equilateral* triangle.

In conclusion, we consider the effect on a polygon of any r-ary S.C., M. Since for an S.C., $\phi(1)=1$ by (3), the determinant of M is by (13)

$$\det M = \phi(\omega)\cdot\phi(\omega^2)\cdot\cdots\cdot\phi(\omega^{n-1}).$$

Hence, if the auxiliary polynomial $\phi(x)$ is prime to x^n-1, then $\det M\neq 0$, and the transformation M is (uniquely) reversible. This means that for every polygon P' there is a (unique) polygon P such that $P'=MP$; accordingly, P' can have no special properties if P is general.

On the other hand, if the greatest common divisor of $\phi(x)$ and x^n-1 is

$$\psi(x) = (x-\omega^{p_1})(x-\omega^{p_2})\cdots(x-\omega^{p_k}),$$

so that

$$\phi(\omega^{p_1})=0,\quad \phi(\omega^{p_2})=0,\cdots,\phi(\omega^{p_k})=0,\quad \phi(\omega^{p})\neq 0 \quad \text{for } p\neq p_1 \text{ or } p_2\cdots \text{ or } p_k, \tag{17}$$

then by (16′) the transformed polygon P' has the special properties (partial regularity of degree k)

[15] Theorem VI is the main one of our cited paper (Theorem A or A′, §7).

$$(18) \qquad R'_{n-p_1} = 0, \qquad R'_{n-p_2} = 0, \qquad \cdots, \qquad R'_{n-p_k} = 0$$

for an arbitrary original polygon P. Further, no other relations $R'_{n-p}=0$ ($p \neq p_1$ or $p_2 \cdots$ or p_k) are satisfied by P' if P remains general (P' has no higher than the kth degree of regularity). This is also seen from (16'), where $\phi(\omega^p) \neq 0$, $R_{n-p} \neq 0$ (since P is general); therefore $R'_{n-p} \neq 0$.

In fact, no relations of any kind besides (18) are satisfied by $P' = MP$ if P remains general. This is because, by the general theory of systems of linear equations, it can be readily shown that if the conditions (17) are satisfied by the coefficients α in (2), then the conditions (18) are sufficient as well as necessary in order that (2) be solvable for the z's in terms of the z''s. This is to say that for *any* polygon P' obeying (18) a polygon P can be found such that $P' = MP$; indeed, the class of such polygons P depends linearly on k complex parameters.

BROOKLYN, N. Y.

AXIOMS FOR MOORE SPACES AND METRIC SPACES[1]

C. W. VICKERY

We shall consider a set of five axioms in terms of the undefined notions of *point* and *region*. It will be shown that these axioms are independent and that they constitute a set of conditions necessary and sufficient for a space to be a complete metric space. It will also be shown that certain subsets of this set of axioms constitute necessary and sufficient conditions for a space to be (1) a metric space, (2) a Moore space, (3) a complete Moore space. Axiom 2 and a more general form of Axiom 1 have been stated by the author in an earlier paper [1]. Following terminology of F. B. Jones [2], a space is said to be a *Moore space* provided conditions (1), (2), and (3) of Axiom 1 (that is, Axiom 1_0) of R. L. Moore's *Foundations of Point Set Theory* [3] are satisfied. A space is said to be a *complete Moore space* provided it satisfies all the conditions of that axiom. Wherever the notion of region is employed, whether as a defined or an undefined notion, it is understood that a necessary and sufficient condition that a point P be a limit point of a point set M is that every region containing P contain a point of M distinct from P. The letter S is used to denote the set of all points.

[1] Presented to the Society, April 20, 1935, under the title *Sets of independent axioms for complete Moore space and complete metric space.*

AXIOM 0. *Every region is a point set.*

AXIOM 1. *There exists a countable family F such that* (1) *every element of F is a collection of regions covering S,* (2) *if R is a region and A and B are points of R, there exists a collection G of F such that if g is a region of G that contains A, $\bar{g}$ is a subset of $(R-B)+A$.*

AXIOM 2. *If P is a point and H and K are regions containing P, there exists a region R containing P which is a subset both of the region H and of K.*

AXIOM 3. *If α is a monotonic descending sequence of closed point sets $A_1, A_2, \cdots$ such that for each n there exists a monotonic descending sequence ρ_n of distinct regions $R_1, R_2, \cdots, R_n$ containing A_n, then there exists a point common to all the elements of α.*

AXIOM 4. *If G is a collection of regions covering S, there exists a collection H of regions covering S such that if h_1 and h_2 are intersecting regions of H, then (h_1+h_2) is a subset of a region of G.*

THEOREM 1. *In order that a space be a Moore space, it is necessary and sufficient that it satisfy Axioms* 0, 1, *and* 2.

The necessity of these conditions is evident. We shall undertake to show their sufficiency without changing the notion of region. Let $H_1, H_2, \cdots$ be a type ω sequence of all the elements of family F postulated by Axiom 1. Let G_1 denote the collection of all regions R such that R is a subset of a region of H_1. Let G_2 denote the collection of all regions R such that R is a subset of a region of H_1 and a region of H_2. For each positive integer n let G_n denote the collection of all regions R such that R is a subset of a region of H_i for each $i \leqq n$. For each n, G_n covers S, by Axiom 2. Furthermore, for each n, G_n contains all the regions of G_{n+1}. The sequence $G_1, G_2, \cdots$ satisfies all the conditions of Axiom 1_0 of R. L. Moore.

As a means to proving the next theorem, we shall prove the following lemma on the basis of Moore's Axioms 0 and 1_0:

LEMMA 1. *If M is a set of points and G is a collection of domains covering M, there exists a collection H of domains covering M such that no domain of H is a subset of another domain of H and such that every domain of H is a subset of some domain of G.*

Suppose that M is a set of points and G a collection of domains covering M. For each positive integer n let T_n denote the set of all points P of M such that some domain of G contains every region of G_n that contains P. Then $M=\sum_{n=1}^{\infty} T_n$. For each positive integer n

let θ_n denote a well-ordered sequence of the points of T_n. Let θ denote the sequence obtained by taking first the elements of θ_1, then the elements of θ_2, and so on. Let $t_{i,\mu}$ denote the first element of θ, where i is the smallest integer such that $t_{i,\mu}$ is an element of θ_i and where μ is an ordinal number denoting the order of $t_{i,\mu}$ in θ_i. (Some sets T_n may be vacuous.) We shall now define a sequence Δ of domains $D_1, D_2, \cdots$. Let D_1 denote the sum of all the regions of G_i that contain $t_{i,\mu}$. Let $t_{j,\nu}$ denote the first point of θ not contained in D_1. Let D_2 denote the sum of all the regions of G_j that contain $t_{j,\nu}$. In general, suppose that Δ_α denotes any *abschnitt* of Δ; then let $t_{k,\xi}$ denote the first point of θ not contained in any domain of Δ_α and let D_α denote the sum of all the regions of G_k that contain $t_{k,\xi}$. Let H denote the collection of all the domains of Δ. Then H has the required properties.

THEOREM 2. *In order that a space be a complete Moore space, it is necessary and sufficient that it satisfy Axioms* 0, 1, 2, *and* 3.

We shall first show the sufficiency of these conditions. Let $H_1, H_2, \cdots$ denote a type ω sequence of the elements of family F of Axiom 1. For each positive integer n let G_n denote the collection of all regions R such that R is a point or a proper subset of a region of H_n and of a region of G_{n-1}. It follows, with the help of Axiom 2, that sequence $G_1, G_2, \cdots$ satisfies conditions (1), (2), and (3) of Moore's Axiom 1. It remains to be shown that it satisfies condition (4). Suppose that $M_1, M_2, \cdots$ is a type ω sequence of nondegenerate closed point sets such that for each n, M_n contains M_{n+1} and is a subset of some region of G_n. Let R_n denote a region of G_n that contains M_n. Then R_n is a proper subset of a region R_{n-1} of G_{n-1}. Similarly R_{n-1} is a proper subset of a region R_{n-2} of G_{n-2}. Thus the conditions of Axioms 3 are satisfied and hence there exists a point common to all the sets $M_1, M_2, \cdots$.

We shall now show the necessity of these conditions by redefining region. Let $G_1, G_2, \cdots$ be a sequence of collections of regions postulated by Moore's Axiom 1. For each n let H_n denote a collection of domains covering S such that no domain of H_n is a subset of another domain of H_n and such that every domain of H_n is a subset of a region of G_n. For each n, H_n exists, by Lemma 1. Let F denote the family of all collections H_n. Let $H=\sum_{n=1}^{\omega} H_n$. If the domains of H are called regions and if nothing else is called a region, then Axioms 0, 1, 2, and 3 are satisfied. (1) Clearly Axioms 0 and 1 are satisfied. (2) We shall show that Axiom 2 is satisfied. Let h and k denote two domains of H having a point P in common. There exists an integer n such that

every region of G_n that contains P is a subset of $h \cdot k$. Let R denote a domain of H_n. Then R is a subset of some region of G_n and hence of $h \cdot k$. (3) We shall now show that Axiom 3 is satisfied. Let α denote a type ω sequence of closed point sets $A_1, A_2, \cdots$, and for each n let ρ_n denote a type n sequence of domains of H, $R_1, R_2, \cdots, R_n$ satisfying the conditions of Axiom 3. Since for each n no domain of H_n is a subset of another domain of H_n, it follows that there exists an $i \geqq n$ such that some domain of ρ_n belongs to H_i and hence is a subset of a region of G_i. It follows that for each n, A_n is a subset of a region of G_n and hence there exists a point common to all the elements of α.

THEOREM 3. *In order that a metric space be complete it is necessary and sufficient that it satisfy Axiom* 3.

This follows immediately with the aid of Theorem 2 and a result of J. H. Roberts [4] to the effect that every metric space that satisfies Axiom 1 of R. L. Moore is complete. In a metric space every interpretation of region that preserves the notion of limit point satisfies Axioms 0, 1, 2, and 4.

THEOREM 4. *In order that a space be metric it is necessary and sufficient that it satisfy Axioms* 0, 1, 2, *and* 4.

We shall first show the sufficiency of these conditions. We have shown that Moore's Axiom 1_0 follows from Axioms 0, 1, and 2. If Axiom 4 be added, it can be shown that the following stronger analogue (due to R. L. Moore) of Moore's Axiom 1_0 follows: "There exists a sequence $G_1, G_2, \cdots$ such that (1) for each n, G_n is a collection of regions covering S, (2) for each n, G_n contains G_{n+1}, (3) if R is a region and A and B are points of R, there exists an integer n such that if h and k are two regions of G_n having a point in common and such that h contains A, then $h+k$ is a subset of $(R-B)+A$." Moore has shown that this proposition is a necessary and sufficient condition for a space to be metric.

We shall show the necessity of these conditions. Suppose that S denotes a space $(\mathcal{D})$. Let all spheroids be called regions. Let collection H_n of family F be the set of all spheroids of radius less than $1/n$. Clearly Axioms 0, 1, and 2 are satisfied. We shall show that Axiom 4 is satisfied. Let G denote a collection of spheroids covering S. For each positive integer n let T_n denote the set of all points P such that there exists a spheroid of G containing the sum of every two-linked chain of spheroids of radius less than $1/n$ that contains P. Then $S = \sum_{n=1}^{\omega} T_n$. Let Q_n denote the collection of all spheroids of radius

less than $1/n$ containing a point of T_n. Let $Q=\sum_{n=1}^{\omega}Q_n$. Then Q is the required collection, for the sum of every two-linked chain of regions of Q is a subset of some region of G.

THEOREM 5. *In order that a space be a complete metric space, it is necessary and sufficient that it satisfy Axioms* 0, 1, 2, 3, *and* 4.

This is an immediate consequence of Theorems 3 and 4.

INDEPENDENCE EXAMPLES

For Axiom 1. Let S be the set of all real numbers between 0 and 1. Let p and q denote two real numbers such that $0<p<q<1$. Let the collection of all regions be the collection of all segments ab such that either (1) $0<a<p$ and $q<b<1$, or (2) $0<a<p$ and $0<b<p$, or (3) $q<a<1$ and $q<b<1$.

For Axiom 2. Let S be the set of all points on the x axis between $(-1, 0)$ and $(+1, 0)$. Let every segment of S not containing O $(0, 0)$ or having O as an end point be taken as a region. Furthermore, let every point set consisting of O together with a segment of S having O as an end point be taken as a region. If n is an odd positive integer, let collection H_n of family F of Axiom 1 be the collection of all regions not containing O and of length less than $1/n$, together with all left-hand regions containing O and of length less than $1/n$. If n is even, we have the same statement except that we substitute right-hand regions containing O for left-hand regions.

For Axiom 3. Let S be the set of all rational points on the x axis. Let the sets of all rational points of all segments be called regions.

For Axiom 4. Let S be the set of all points on or above the x axis. Let regions be the interiors of all circles lying wholly above the x axis, together with all point sets Q such that Q is the interior of a circle tangent to the x axis plus the point of tangency. (Example due to R. L. Moore.)

REFERENCES

1. C. W. Vickery, *Spaces in which there exist uncountable convergent sequences of points*, Tôhoku Mathematical Journal, vol. 40 (1934), pp. 1–26.

2. F. B. Jones, *Concerning normal and completely normal spaces*, this Bulletin, vol. 43 (1937), pp. 671–677.

3. R. L. Moore, *Foundations of Point Set Theory*, American Mathematical Society Colloquium Publications, vol. 13, New York, 1932.

4. J. H. Roberts, *A property related to completeness*, this Bulletin, vol. 38 (1932), pp. 835–838.

AUSTIN, TEXAS

THE AUTOMORPHISMS OF THE SYMMETRIC GROUP

IRVING E. SEGAL

The purpose of this note is to give a proof of the following well known theorem. *The group of automorphisms of the symmetric group S_n on n letters is isomorphic with S_n, except when $n=6$.* The proofs of this in the literature are complicated[1] and involve the use of lemmas whose relevance is not plain.

Let A be an automorphism of S_n. Then it is clear that A takes a class of similar elements into a class of similar elements, and that it takes an element of order m into an element with the same order. Hence suppose $A(1r)=t_1(r)\cdot t_2(r)\cdot\cdots\cdot t_k(r)$ $(k\geqq 1)$, where the $t_i(r)$ are disjoint transpositions. A simple calculation shows that there are $n(n-1)/2$ elements similar to $(1r)$, and that there are $n!/2^k k!(n-2k)!$ elements similar to $t_1(r)\cdot t_2(r)\cdot\cdots\cdot t_k(r)$. Hence

$$\frac{n(n-1)}{2}=\frac{n!}{2^k k!(n-2k)!}.$$

If $n\neq 6$ this equation is satisfied for no k $(k\geqq 1)$ except $k=1$.

Suppose now that $n\neq 6$. Then $A(1r)=(a_r b_r)$ say. If $r\neq 2$, $(12)(1r)=(12r)$ (multiplying from right to left), and evidently, $A(12r)=(a_2b_2)(a_rb_r)$. Since $(12r)$ has the order 3, so has $(a_2b_2)(a_rb_r)$ and the transpositions (a_2b_2) and (a_rb_r) must have a letter in common. Then it is no loss to assume $a_2=a_r$ or $b_2=b_r$. However, if $a_2=a_r$ and $b_2=b_s$ $(r\neq 2,\ s\neq 2)$, then $r\neq s$ and $A(12r)=A(12)\cdot A(1r)=(a_2b_2)(a_2b_r)=(b_ra_2b_2)$. Similarly $A(12s)=(a_sb_2a_2)$. Hence $A((12r)\cdot(12s))=A(12r)\cdot A(12s)=(b_ra_2b_2)(a_sb_2a_2)=(b_ra_sb_2)$ which is of order 3, while $(12r)\cdot(12s)=(1s)(2r)$, which is of order 2. Hence one must have $a_2=a_r$ for all r or $b_2=b_r$ for all r; of course one can let $a_2=a_r$ $(r=2, 3, \cdots, n)$. Then $A(1r)=(a_2b_r)$. Hence A is precisely the automorphism A defined by $Ax=t^{-1}xt$, where

$$t=\begin{pmatrix}1 & 2 & \cdots & r & \cdots & n\\ a_2 & b_2 & \cdots & b_r & \cdots & b_n\end{pmatrix}.$$

For $Ax=t^{-1}xt$ when $x=(1r)$, and the elements $\{(1r)\}$ $(r=2, 3, \cdots, n)$ generate S_n.

YALE UNIVERSITY

[1] The first proof is by O. Hölder, Mathematische Annalen, vol. 46 (1895), especially pp. 340–345.

A CORRECTION TO "A NOTE ON LINEAR FUNCTIONALS"[1]

R. P. BOAS, JR., AND J. W. TUKEY

R. S. Phillips has called our attention to an error in our paper *A note on linear functionals*. On page 526, we have misquoted a theorem of Lebesgue's: the statement in the last display on that page is incorrect. It is, in fact, contradicted by the Riemann-Lebesgue theorem whenever the functions $x_n(t)$ are the elements of a uniformly bounded orthonormal set. Fortunately, however, the error does not affect the validity of any of our results. The correct consequence of Lebesgue's theorem is that

$$\sup_{0\leqq n<\infty} \sup_{0\leqq t\leqq 1}{}^{0} \; | x_n(t) | < \infty ; \tag{1}$$

that is, that $\sup_{0\leqq n<\infty}\|x_n\|_B<\infty$. From this it still follows that any linear functional on B is a linear functional on R; and we used our incorrect statement only to deduce this. This consequence is true in virtue of the following simple lemma.

LEMMA. *If a set* $\{x\}$ *forms a normed vector space under two norms,* $\|x\|$ *and* $\|x\|_B$, *and if* $\lim_{n\to\infty}\|x_n\|=0$ *implies that* $\sup_{0\leqq n<\infty}\|x_n\|_B<\infty$, *then any distributive functional continuous with respect to the second norm is also continuous with respect to the first norm.*

PROOF. Let f be a distributive functional, continuous with respect to the norm $\|\cdots\|_B$, so that for some number H,

$$| f(x) | \leqq H\|x\|_B \tag{2}$$

for every x. Suppose that f is not continuous with respect to the norm $\|\cdots\|$; then, as is well known (cf. S. Banach, *Théorie des Opérations Linéaires*, 1932, p. 55) there exist elements y_n such that $\|y_n\|=1$, $|f(y_n)|>n$. The elements $z_n=n^{-1/2}y_n$ have the properties

$$\|z_n\| \to 0, \tag{3}$$

$$| f(z_n) | > n^{1/2}. \tag{4}$$

By hypothesis, (3) implies that $\|z_n\|_B<K$, $n=0, 1, 2, \cdots$, for some finite K. Then, by (2), $|f(z_n)|\leqq HK$, contradicting (4) for large n.

DUKE UNIVERSITY AND
PRINCETON UNIVERSITY

[1] This Bulletin, vol. 44 (1938), pp. 523–528.